INTERNATIONAL STANDARD

ISO 12199

First edition
2000-08-01

Alphabetical ordering of multilingual terminological and lexicographical data represented in the Latin alphabet

Mise en ordre alphabétique des données lexicographiques et terminologiques multilingues représentées dans l'alphabet latin

Reference number
ISO 12199:2000(E)

© ISO 2000

PDF disclaimer

This PDF file may contain embedded typefaces. In accordance with Adobe's licensing policy, this file may be printed or viewed but shall not be edited unless the typefaces which are embedded are licensed to and installed on the computer performing the editing. In downloading this file, parties accept therein the responsibility of not infringing Adobe's licensing policy. The ISO Central Secretariat accepts no liability in this area.

Adobe is a trademark of Adobe Systems Incorporated.

Details of the software products used to create this PDF file can be found in the General Info relative to the file; the PDF-creation parameters were optimized for printing. Every care has been taken to ensure that the file is suitable for use by ISO member bodies. In the unlikely event that a problem relating to it is found, please inform the Central Secretariat at the address given below.

Contents   Page

ISO 12199:2000(E)

Foreword

ISO (the International Organization for Standardization) is a worldwide federation of national standards bodies (ISO member bodies). The work of preparing International Standards is normally carried out through ISO technical committees. Each member body interested in a subject for which a technical committee has been established has the right to be represented on that committee. International organizations, governmental and non-governmental, in liaison with ISO, also take part in the work. ISO collaborates closely with the International Electrotechnical Commission (IEC) on all matters of electrotechnical standardization.

International Standards are drafted in accordance with the rules given in the ISO/IEC Directives, Part 3.

Draft International Standards adopted by the technical committees are circulated to the member bodies for voting. Publication as an International Standard requires approval by at least 75 % of the member bodies casting a vote.

Attention is drawn to the possibility that some of the elements of this International Standard may be the subject of patent rights. ISO shall not be held responsible for identifying any or all such patent rights.

International Standard ISO 12199 was prepared by Technical Committee ISO/TC 37, *Terminology (principles and coordination)*, Subcommittee SC 2, *Layout of vocabularies.*

It complements other International Standards prepared by ISO/TC 37, such as ISO 10241:1992, *International terminology standards — Preparation and layout* and ISO 12200:1999, *Computer applications in terminology — Machine-readable terminology interchange format (MARTIF) — Negotiated interchange.*

Annexes A and G form a normative part of this International Standard. Annexes B to F are for information only.

Introduction

In the development of international terminologies, both in printed form and in databases, it is essential to have uniform and internationally recognized rules for the alphabetical ordering of terminological and lexicographical data, to make these terminologies more easily accessible for the users. In addition, it will facilitate the interchange of terminological and lexicographical data.

INTERNATIONAL STANDARD
ISO 12199:2000(E)

Alphabetical ordering of multilingual terminological and lexicographical data represented in the Latin alphabet

1 Scope

This International Standard specifies the sequence of characters to be used in the alphabetical ordering of multilingual terminological and lexicographical data (terms, term elements, or words) represented in the Latin alphabet. Character sets of languages represented in the Latin alphabet are taken into account insofar as terminological or lexicographical data have been recorded. Character sets used in internationally standardized transliteration into Latin script are also taken into account.

The sequence of alphabetical characters given is intended for multilingual purposes only and is not intended to affect the alphabetical order of any specific language.

The main part of this International Standard specifies letter-by-letter ordering of character strings. Normative annex A treats word-by-word ordering, which is a widely used alternative to this system.

Informative annex B gives two additional rules that may be useful for lexicographical and terminological ordering.

Informative annex C gives ordering rules for chemical names.

Informative annex D lists the character repertoire of the Latin alphabet.

Informative annex E lists languages using the Latin alphabet.

Informative annex F gives alphabetical sequences derived from the sequence specified in this International Standard for a number of languages that use the Latin alphabet.

Normative annex G gives a formal description of the rules laid down in the main part of this International Standard conforming with ISO/IEC 14651.

2 Normative references

The following normative documents contain provisions which, through reference in this text, constitute provisions of this International Standard. For dated references, subsequent amendments to, or revisions of, any of these publications do not apply. However, parties to agreements based on this International Standard are encouraged to investigate the possibility of applying the most recent editions of the normative documents indicated below. For undated references, the latest edition of the normative document referred to applies. Members of ISO and IEC maintain registers of currently valid International Standards.

ISO 1087:1990, *Terminology — Vocabulary.*

ISO 1087-1:—[1], *Terminology work — Vocabulary — Part 1: Theory and application.*

ISO 1087-2:2000, *Terminology work — Vocabulary — Part 2: Computer applications.*

ISO/IEC 10646-1:1993, *Information technology — Universal Multiple-Octet Coded Character Set (UCS) — Part 1: Architecture and Basic Multilingual Plane.*

ISO/IEC 14651:—[1], *Information technology — International string ordering — Method for comparing character strings and description of a default tailorable ordering.*

---

1) To be published.

3 Terms and definitions

For definitions of terminological concepts, see ISO 1087, ISO 1087-1 and ISO 1087-2.

For the purpose of this International Standard, the following terms and definitions apply.

3.1
character
member of a set of elements used for the organization, control or representation of data

3.2
letter
character used for writing natural language, often representing a sound in the language

3.3
digit
character used to represent the numeric value, or part thereof, of a number

3.4
special character
character that is not a letter nor a digit

EXAMPLE   The space character is a special character.

3.5
ligature
character resulting from the joining of two or more letters

NOTE   The resulting character is, in some cases, considered a separate letter.

3.6
polygraph
two or more consecutive letters that are regarded as one letter for some purpose

NOTE   A polygraph consisting of two or three letters may be referred to as a digraph or a trigraph respectively.

3.7
diacritical mark
character that is not a letter and is placed over, under, or through a letter or a combination of letters

3.8
ordering
act of bringing strings of characters into a well-defined sequence according to a string comparison specification

4 Preparatory procedures

In the process of alphabetical ordering, character strings are compared according to a set of rules. This International Standard specifies the set of rules to be used for the ordering, but does not address the means of selection of relevant character strings, nor any modification of the strings that may be needed for a given purpose. Consequently, certain preparatory procedures may be needed before applying the ordering rules. Depending on the needs in each individual case

— the relevant character strings may have to be selected, e.g. relevant terms may have to be extracted from a corpus,

— the character strings may have to be modified, e.g. sentence-initial uppercase letters may have to be changed to lowercase letters, plural form of words may have to be changed to singular form, or

— leading zeroes or spaces may be added e.g. in lists containing numerals.

Polygraphs are treated as sequences of separate letters.

An application may arrange information into several ordering fields, and determine ranking order with several separate and independent comparisons. This International Standard only defines a single comparison for one such field, where the field is a character-string field.

Only the characters that appear in the string and their arrangement are taken into account. Apart from the ordering rules and passes, no other knowledge about the words in the character string is used. For example, dictionary information or rules about language syntax, phonetics and semantics are not used.

5 First ordering level

5.1 First-ordering-level values

When comparing strings to be ordered, the first-ordering-level values of the strings shall be considered first. The subsequent ordering-level values need to be considered only if two or more strings have identical first-ordering-level values.

For multilingual ordering, the following rules shall be applied (see annex A for word-by-word ordering):

5.2 First-ordering-level sequence

Digits and letters have the following ordering values:

a) **Digits:** 0 1 2 3 4 5 6 7 8 9

NOTE 1   Sequences of digits will be ordered from left to right as written, thus generating the following order, e.g.: 1 10 100 11 110 111 12 19 190 2 21 3.

NOTE 2   Leading zeroes may be inserted as a preparatory procedure, e.g. to generate the following order: 0001 0002 0003 0010 0011 0012 0019 0021 0100 0110 0111 0190.

b)   **Basic letters of the Latin alphabet:**

a A	b B	c C	d D	e E	f F	g G	h H	i I	j J	k K	l L	m M	n N
o O	p P	q Q	r R	s S	t T	u U	v V	w W	x X	y Y	z Z	þ Þ	

NOTE 1   This order has been established for use in multilingual environments so as to conflict with as few individual languages as possible. See informative annex F for examples of deviations from this sequence in some languages.

Uppercase and lowercase letters shall be treated as equivalent (see clause 7). Letters of the Latin alphabet with diacritical marks shall be treated as equivalent to the corresponding basic Latin letters (see clause 6). Special letters of the Latin alphabet shall be treated as equivalent to basic Latin letters according to Table 1 in 5.3 (see clause 6).

The Turkish language distinguishes ı/I from i/İ, while other languages have the pair i/I only. To order multilingual data including Turkish text, the i/I pair shall be expanded as follows:

1:	ı/I	U0131/U0049	LATIN LETTER DOTLESS I (Turkish)
2:	i/I	U0069/U0049	LATIN LETTER I (non-Turkish)
3:	i/İ	U0069/U0130	LATIN LETTER I WITH DOT ABOVE (Turkish)

It should also be noted that, for example, í (U00ED LATIN SMALL LETTER I WITH ACUTE) in normal print is represented as LATIN SMALL LETTER DOTLESS I WITH ACUTE. For the purpose of ordering, however, it shall be treated as equivalent to i (U0069 LATIN SMALL LETTER I) on the first ordering level.

NOTE 2   Throughout this International Standard, characters are referenced as UXXXX, where X is any hexadecimal digit and refers to the position of the character in ISO/IEC 10646-1. Character names are given as in ISO/IEC 10646-1. Most names of Latin letters start with "LATIN SMALL LETTER ..." and "LATIN CAPITAL LETTER ...". When referring to both lowercase and uppercase letter, the name "LATIN LETTER ..." is used. When there is no danger of misinterpretation, the words "LATIN LETTER" are sometimes omitted.

ISO 12199:2000(E)

c) Letters of other alphabets

Letters of other alphabets follow in the sequences established for each alphabet. The order of non-Latin alphabets shall be: the Greek alphabet, the Cyrillic alphabet, other alphabets.

NOTE   It is outside the scope of this International Standard to establish the sequences for alphabets other than the Latin alphabet. The Greek alphabet has the following sequence of letters:

α A   β B   γ Γ   δ Δ   ε E   ζ Z   η H   θ Θ   ι I   κ K   λ Λ   μ M   ν N   ξ Ξ
ο O   π Π   ρ P   σ Σ   τ T   υ Y   φ Φ   χ X   ψ Ψ   ω Ω.

All other characters, e.g. punctuation marks, shall be ignored. See clause 8.

5.3   Equivalence between special Latin letters and basic letters

Special Latin letters shall be treated as equivalent to basic letters of the Latin alphabet according to Table 1. Uppercase and lowercase letters shall be treated as equivalent.

Table 1 — Equivalence between special Latin letters and basic letters

Position	Character name in ISO/IEC 10646-1	Character position for lowercase / uppercase in ISO/IEC 10646-1		Equivalent to
01	LATIN LETTER AE	U00E6	U00C6	ae
02	LATIN LETTER B WITH HOOK	U0253	U0181	b
03	LATIN LETTER C WITH HOOK	U0188	U0187	c
04	LATIN LETTER D WITH STROKE	U0111	U0110	d
05	LATIN LETTER D WITH HOOK	U0257	U018A	d
06	LATIN LETTER ETH	U00F0	U00D0	d
07	LATIN LETTER G WITH HOOK	U0260	U0193	g
08	LATIN LETTER H WITH STROKE	U0127	U0126	h
09	LATIN LETTER K WITH HOOK	U0199	U0198	k
10	LATIN SMALL LETTER KRA	U0138	a	k
11	LATIN LETTER L WITH STROKE	U0142	U0141	l
12	LATIN LETTER ENG	U014B	U014A	n
13	LATIN LETTER O WITH STROKE	U00F8	U00D8	o
14	LATIN LIGATURE OE	U0153	U0152	oe
15	LATIN SMALL LETTER SHARP S	U00DF	a	ss
16	LATIN LETTER T WITH STROKE	U0167	U0166	t

a   No corresponding uppercase letter.

6   Second ordering level

6.1   Second-ordering-level values

If the comparison of two strings results in identical first-ordering-level values, second-ordering-level values shall be applied according to 6.2.

The rule shall be applied from left to right.

6.2   Special Latin letters and letters with diacritical marks

Special Latin letters, that have been treated as equivalent to basic Latin letters according to Table 1, shall be ordered according to the order in Table 1.

Diacritical marks shall be ordered according to Table 2.

NOTE   This order has been established for multilingual environments so as to be in conflict with as few individual languages as possible. See informative annex F for examples of deviations from this sequence in some languages.

Table 2 — Ordering of diacritical marks

Position	Name	Position for combining diacritical mark in ISO/IEC 10646-1
0000	none	
0100	ACUTE ACCENT	U0301
0200	GRAVE ACCENT	U0300
0300	BREVE	U0306
0301	BREVE AND ACUTE	—
0302	BREVE AND GRAVE	—
0310	BREVE AND HOOK ABOVE	—
0311	BREVE AND TILDE	—
0313	BREVE AND DOT BELOW	—
0315	BREVE AND COMMA BELOW	—
0400	CIRCUMFLEX ACCENT	U0302
0401	CIRCUMFLEX AND ACUTE	—
0402	CIRCUMFLEX AND GRAVE	—
0410	CIRCUMFLEX AND HOOK ABOVE	—
0411	CIRCUMFLEX AND TILDE	—
0413	CIRCUMFLEX AND DOT BELOW	—
0500	CIRCUMFLEX ACCENT BELOW	U032D
0600	CARON	U030C
0614	CARON AND CEDILLA	—
0700	RING ABOVE	U030A
0701	RING ABOVE AND ACUTE	—
0800	DIAERESIS	U0308
0813	DIAERESIS AND DOT BELOW	—
0817	DIAERESIS AND MACRON	—
0900	DOUBLE ACUTE ACCENT	U030B
1000	HOOK ABOVE	U0309
1100	TILDE	U0303
1200	DOT ABOVE	U0307
1300	DOT BELOW	U0323
1400	CEDILLA	U0327
1500	COMMA ABOVE/BELOW	U0313 and U0326 [a]
1600	OGONEK	U0328
1700	MACRON	U0304
1713	MACRON AND DOT BELOW	—
1800	MACRON BELOW	U0331
1900	PRECEDED BY APOSTROPHE	—
2000	FOLLOWED BY APOSTROPHE	—
2100	HORN	U031B
2101	HORN AND ACUTE	—
2102	HORN AND GRAVE	—
2110	HORN AND HOOK ABOVE	—
2111	HORN AND TILDE	—
2113	HORN AND DOT BELOW	—

[a] The position of combining comma above and below the base character.

7 Third ordering level

7.1 Third-ordering-level values

If the comparison of two strings results in identical first- and second-ordering-level values, third-ordering-level values shall be applied according to 7.2.

The rule shall be applied from left to right.

7.2 Ordering according to capitalization

A lowercase letter shall be ordered before the corresponding uppercase letter. [See 5.2, item b), first paragraph after note 1.]

NOTE   The terms "lowercase letter" and "uppercase letter" are used for members of the sets "a b c ..." and "A B C ...", respectively. In character names, the naming conventions of ISO/IEC 10646-1 are used. ISO/IEC 10646-1 uses "LATIN SMALL LETTER" and "LATIN CAPITAL LETTER", respectively.

8 Fourth ordering level

8.1 Fourth-ordering-level values

If the comparison of two strings results in identical first-, second- and third-ordering-level values, fourth-ordering-level values shall be applied according to 8.2.

The rule shall be applied from left to right.

8.2 Ordering according to special characters

Special characters are ordered according to the sequence of the default template of ISO/IEC 14651. For most special characters, this is the order in which they are listed in ISO/IEC 10646-1.

NOTE   In word-by-word ordering (see normative annex A), the space character and possibly other special characters may have special functions as key separators.

Annex A
(normative)

Word-by-word ordering

A.1 Principles of word-by-word ordering

As noted in the scope, this International Standard specifies the letter-by-letter ordering of character strings. Word-by-word ordering is a widely used alternative to this system. Table A.1 illustrates the difference between letter-by-letter ordering and word-by-word ordering.

Table A.1 — Letter-by-letter and word-by-word ordering

Letter-by-letter ordering	Word-by-word ordering
ad	ad
adhesive	ad hoc
ad hoc	ad infinitum
adieu	adhesive
ad infinitum	adieu
adipose	adipose

A.2 Multiple-key ordering

Single-key ordering is described in the main body of this International Standard. In multiple-key ordering, all the ordering rules are applied to one key before they are applied to the next, until all the keys have been considered or a unique sequence has been established.

NOTE   One typical example of multiple-key ordering is a list of delegates to a meeting, where the first key may be the country names, the second key may be the delegates' last names, and the third key may be the delegates' first names. In this example, if a country has one delegate only, the second key (last names) will not be considered.

A.3 Word-by-word ordering as multiple-key ordering

In word-by-word ordering, space characters, and possibly also by definition other characters, are key separators. The key-separator characters function as key separators only, and they have no position in the ordering sequence.

When the character string has been divided into a sequence of keys, the ordering rules of the main body of this International Standard are invoked for one key at a time.

NOTE 1   In addition to the space characters, some or all punctuation marks may be defined as key separators. It may also be useful to define some space characters as key separators, while other space characters remain special characters within a key. The choices will depend on language(s) and type of strings to be ordered.

NOTE 2   If space characters and hyphens are defined as key separators, the title of this clause would be split into the following keys: <A.3> <Word> <by> <word> <ordering> <as> <multiple> <key> <ordering>, where each key is contained within < and >, and the spaces are added for increased readability.

A.4 Simple word-by-word ordering

If the text to be ordered using word-by-word ordering contains very few special Latin letters and diacritical marks, the following extension to the rules in the main body of this International Standard will produce the same or nearly the same output as the rules described in clause A.3.

On the first ordering level (see 5.2), the **space character** is added as the first item. Items 1, 2, and 3 in 5.2 then become items 2, 3, and 4. The space character is not treated as a special character on the fourth ordering level (clause 8).

NOTE   Depending on the language(s) and type of strings to be ordered, it may be useful to treat even other special characters (e.g. hyphens) in the same way as the space character.

Annex B
(informative)

Special rules for lexicographical and terminological ordering

B.1 Background

For lexicographical and terminological applications, it may sometimes be desirable to add additional ordering criteria to the rules that are described in the main body of this International Standard.

The features that are described in this annex cannot easily be described in the formalism given in ISO/IEC 14651.

B.2 Position relative to baseline

It may be desirable to distinguish, for example, m2, m^2, m_2 for ordering purposes. If this is deemed necessary, it is recommended that this be done on the third ordering level (see clause 7) combined with capitalization.

The ordering value of any given character based on its position relative to the baseline may be determined according to Table B.1.

Table B.1 — Position relative to baseline

1	character(s) on baseline
2	character(s) above baseline, superscript character(s)
3	character(s) below baseline, subscript character(s)

B.3 Ordering according to styles

If ordering by the first through fourth ordering level does not produce a unique sequence, typographical styles may be taken into consideration as a fifth ordering level.

Styles may be ordered according to Table B.2.

Table B.2 — Order of styles

Position	Style name	Example
1	roman	abcdefghij
2	boldface	**abcdefghij**
3	italic	*abcdefghij*
4	boldface-italic	***abcdefghij***
5	others	

Annex C
(informative)

Ordering rules for chemical names

C.1 Background

There are no universally accepted ordering rules for chemical names. The ordering rules of the main body of this International Standard may be used, if so desired with the extension of the word-by-word ordering rules described in annex A.

However, some indexes and databases, in particular at the Chemical Abstracts Services (CAS), use a specially designed multiple-key ordering system. The main features of this system are outlined below.[2]

C.2 Division into three keys

C.2.1 Parent name

The first key consists of the parent name, which normally will be all roman letters and space characters, whether or not interrupted by italic letters, Greek letters, digits or special characters (e.g. punctuation).

C.2.2 Initial locants

The second key consists of initial locants, being all characters before the first roman letter.

C.2.3 Other locants

The third key consists of all non-initial locants, being all remaining characters.

NOTE   The name "2-Butanone-1,1,1-d_3, 3,3-dimethyl" is divided into three keys as follows: <Butanone dimethyl> <2-> <-1,1,1-d_3, 3,3->

C.3 Ordering rules within each key

The first key is ordered according to the rules of the main body of this International Standard.

In the second and third keys, the following order is used:

— letters of the Latin alphabet (which will be in italic), in the order specified in 5.2, item b);

— letter of the Greek alphabet, in the order given in 5.2, item c);

— numerals, in the order of the numeric value.

C.4 Output

Table C.1 shows ordered output from the rules that are described in this annex compared with output from the rules of the main body of this International Standard.

---

2) For further details, please consult Chemical Abstracts Services (CAS), P.O. Box 3012, Columbus, Ohio 43210, USA.

Table C.1 — Sample output

Ordered according to annex C	Ordered according to general rules
Bromine fluoride (BrF$_3$)	1-Butanone
Bromine fluoride (BrF$_5$)	1-Butanone, 1-phenyl-
2-Butanol	2-Butanol
2-Butanol, (R)-	2-Butanol, 1-chloro-
2-Butanol, (S)-	2-Butanol, 4-(trimethylstannyl)-
2-Butanol, sodium salt, (S)-	2-Butanol, (R)-
2-Butanol, 1-chloro-	2-Butanol, (S)-
2-Butanol, 4-(trimethylstannyl)-	2-Butanol, sodium salt, (S)-
1-Butanone	2-Butanone
1-Butanone, 1-phenyl-	2-Butanone, 1-(dimethylamino)-3,3-dimethyl-
2-Butanone	2-Butanone-1,1,1-d_3
2-Butanone, O-methyloxime	2-Butanone-1,1,1-d_3, 3,3-dimethyl-
2-Butanone, oxime	2-Butanone, 3-(4-acetylphenyl)-
2-Butanone, polymer with formaldehyde	2-Butanone, 3-ethoxy-1,1-dihydroxy-
2-Butanone, 3-(4-acetylphenyl)-	2-Butanone, O-methyloxime
2-Butanone, 1-(dimethylamino)-3,3-dimethyl-	2-Butanone, oxime
2-Butanone, 3-ethoxy-1,1-dihydroxy-	2-Butanone, polymer with formaldehyde
2-Butanone-1,1,1-d_3	Bromine fluoride (BrF$_3$)
2-Butanone-1,1,1-d_3, 3,3-dimethyl-	Bromine fluoride (BrF$_5$)
Butanoyl chloride	Butanoyl chloride

Annex D
(informative)

Character repertoire of the Latin alphabet

Table D.1 lists the character repertoire of the Latin alphabet. The languages listed in annex E have been taken into account if reliable information is available. Characters that are exclusive to the International Phonetic Alphabet have not been included.

NOTE   The names used in ISO/IEC 10646-1 are used in the "Name" column of Table D.1. The full names of the letters are "LATIN SMALL LETTER ..." and "LATIN CAPITAL LETTER ..." for the lowercase and uppercase letters respectively. In the "Type" column of Table D.1, b = basic Latin letter; d = Latin letter with diacritical mark; s = special Latin letter. In the column "languages used", + indicates that the letter is used in most or all languages that use the Latin alphabet. For the language symbols used in Table D.1, see annex E and annex F. The language symbols in square brackets refer to transliteration systems; see Table F.2.

Table D.1 — Character repertoire

Name	Type	Position for lowercase/uppercase in ISO/IEC 10646-1		Languages used
LATIN LETTER A	b	U0061	U0041	+
WITH ACUTE	d	U00E1	U00C1	af ca cs cy da es fo fur ga gd gl hu is kl nl pt qal sk smi ss vi [Cyr] [ar]
WITH GRAVE	d	U00E0	U00C0	ca cy de fr fur fy gd it nl no pt qal rm vi [Cyr]
WITH BREVE	d	U0103	U0102	mo ro vi [Cyr]
WITH BREVE AND ACUTE	d	U1EAF	U1EAE	vi
WITH BREVE AND GRAVE	d	U1EB1	U1EB0	vi
WITH BREVE AND HOOK ABOVE	d	U1EB3	U1EB2	vi
WITH BREVE AND TILDE	d	U1EB5	U1EB4	vi
WITH BREVE AND DOT BELOW	d	U1EB7	U1EB6	vi
WITH CIRCUMFLEX	d	U00E2	U00C2	br cy de fr fur fy kl mo pt qal rm ro smi vi [Cyr] [ar]
WITH CIRCUMFLEX AND ACUTE	d	U1EA5	U1EA4	vi
WITH CIRCUMFLEX AND GRAVE	d	U1EA7	U1EA6	vi
WITH CIRCUMFLEX AND HOOK ABOVE	d	U1EA9	U1EA8	vi
WITH CIRCUMFLEX AND TILDE	d	U1EAB	U1EAA	vi
WITH CIRCUMFLEX AND DOT BELOW	d	U1EAD	U1EAC	vi
WITH CARON	d	U01CE	U01CD	[Cyr]
WITH RING ABOVE	d	U00E5	U00C5	da kl no smi sv [Cyr]
WITH RING ABOVE AND ACUTE	d	U01FB	U01FA	
WITH DIAERESIS	d	U00E4	U00C4	cy de et fi fy lb nl sk smi sv tr [Cyr]
WITH DIAERESIS AND DOT BELOW	d	–	–	[Cyr]
WITH DIAERESIS AND MACRON	d	U01DF	U01DE	
WITH DOUBLE ACUTE	d	–	–	[Cyr]
WITH HOOK ABOVE	d	U1EA3	U1EA2	vi
WITH TILDE	d	U00E3	U00C3	kl pt vi

Name	Type	Position for lowercase/uppercase in ISO/IEC 10646-1		Languages used
WITH DOT BELOW	d	U1EA1	U1EA0	vi
WITH OGONEK	d	U0105	U0104	lt pl
WITH MACRON	d	U0101	U0100	lv [Cyr] [ar]
LATIN LETTER AE	s	U00E6	U00C6	da fo fr is kl no smi [Cyr]
WITH ACUTE	s d	U01FD	U01FC	
WITH MACRON	s d	U01E3	U01E2	
LATIN LETTER B	b	U0062	U0042	+
WITH DOT ABOVE	d	U1E03	U1E02	[he]
WITH DOT BELOW	d	U1E05	U1E04	
WITH HOOK	s	U0253	U0181	ha
LATIN LETTER C	b	U0063	U0043	+
WITH ACUTE	d	U0107	U0106	hr pl [Cyr]
WITH GRAVE	d	–	–	[Cyr]
WITH BREVE	d	–	–	[Cyr]
WITH BREVE AND COMMA BELOW	d	–	–	[Cyr]
WITH CIRCUMFLEX	d	U0109	U0108	eo [Cyr]
WITH CARON	d	U010D	U010C	cs hr lt lv sk sl smi [Cyr]
WITH DIAERESIS	d	–	–	[Cyr]
WITH DOT ABOVE	d	U010B	U010A	mt
WITH DOT BELOW	d	–	–	[Cyr]
WITH CEDILLA	d	U00E7	U00C7	ca fr oc pt sq tr [Cyr]
WITH MACRON	d	–	–	[Cyr]
WITH HOOK	s	U0188	U0187	
LATIN LETTER D	b	U0064	U0044	+
WITH CIRCUMFLEX	d	–	–	[Cyr]
WITH CIRCUMFLEX BELOW	d	U1E13	U1E12	hz ve
WITH CARON	d	U010F		(cs) (sk)
WITH CARON	d		U010E	cs sk
WITH DOT ABOVE	d	U1E0B	U1E0A	
WITH DOT BELOW	d	U1E0D	U1E0C	[ar]
WITH LINE BELOW	d	U1E0F	U1E0E	[ar]
FOLLOWED BY APOSTROPHE	d	–		cs sk
WITH STROKE [a]	s	U0111	U0110	hr smi vi [Cyr]
WITH HOOK	s	U0257	U018A	ha
LATIN LETTER ETH [a]	s	U00F0	U00D0	fo is
LATIN LETTER E	b	U0065	U0045	+
WITH ACUTE	d	U00E9	U00C9	af ca cs cy da de es fr fy ga gd gl hu is it kl lb nl no pt qal sk sl ss sv vi
WITH GRAVE	d	U00E8	U00C8	af ca cy de fr fur gd it nl no pt qal rm vi [Cyr]
WITH BREVE	d	U0115	U0114	[Cyr]
WITH CIRCUMFLEX	d	U00EA	U00CA	af br cy de fr fy nl no nso pt qal rm sl tn vi [Cyr]
WITH CIRCUMFLEX AND ACUTE	d	U1EBF	U1EBE	vi
WITH CIRCUMFLEX AND GRAVE	d	U1EC1	U1EC0	vi
WITH CIRCUMFLEX AND HOOK ABOVE	d	U1EC3	U1EC2	vi
WITH CIRCUMFLEX AND TILDE	d	U1EC5	U1EC4	vi

Name	Type	Position for lowercase/uppercase in ISO/IEC 10646-1		Languages used
WITH CIRCUMFLEX AND DOT BELOW	d	U1EC7	U1EC6	vi
WITH CIRCUMFLEX BELOW	d	U1E19	U1E18	
WITH CARON	d	U011B	U011A	cs [Cyr]
WITH DIAERESIS	d	U00EB	U00CB	af cy de fr fy lb nl sq [Cyr]
WITH HOOK ABOVE	d	U1EBB	U1EBA	vi
WITH TILDE	d	U1EBD	U1EBC	vi
WITH DOT ABOVE	d	U0117	U0116	lt
WITH DOT BELOW	d	U1EB9	U1EB8	vi
WITH OGONEK	d	U0119	U0118	lt pl
WITH MACRON	d	U0113	U0112	lv
LATIN LETTER F	b	U0066	U0046	+
WITH GRAVE	d	–	–	[Cyr]
LATIN LETTER G	b	U0067	U0047	+
WITH ACUTE	d	U01F5	U01F4	[Cyr]
WITH GRAVE	d	–	–	[Cyr]
WITH BREVE	d	U011F	U011E	tr [Cyr]
WITH CIRCUMFLEX	d	U011D	U011C	eo
WITH CARON	d	U01E7	U01E6	[ar]
WITH DOT ABOVE	d	U0121	U0120	mt [Cyr] [ar]
WITH COMMA BELOW/ABOVE [b]	d	–	–	lv
WITH HOOK	s	U0260	U0193	
LATIN LETTER H	b	U0068	U0048	+
WITH CIRCUMFLEX	d	U0125	U0124	eo
WITH DOT ABOVE	d	U1E23	U1E22	[he]
WITH DOT BELOW	d	U1E25	U1E24	[Cyr] [ar] [he]
WITH CEDILLA	d	U1E29	U1E28	[Cyr]
WITH LINE BELOW	d	U1E96	–	[ar]
WITH STROKE	s	U0127	U0126	mt
LATIN LETTER I	b	U0069	U0049	+
LATIN SMALL LETTER DOTLESS I	b	U0131		tr
LATIN CAPITAL LETTER I WITH DOT ABOVE	b		U0130	tr
LATIN LETTER I WITH ACUTE	d	U00ED	U00CD	af ca cs cy da es fo ga gd gl hu is it kl nl pt qal sk vi [Cyr] [ar]
WITH GRAVE	d	U00EC	U00CC	cy fur gd it qal vi [Cyr]
WITH BREVE	d	U012D	U012C	
WITH CIRCUMFLEX	d	U00EE	U00CE	af cy fr it kl mo qal ro [Cyr]
WITH CARON	d	U01D0	U01CF	[Cyr]
WITH DIAERESIS	d	U00EF	U00CF	af ca cy de fr fy it nl oc [Cyr]
WITH HOOK ABOVE	d	U1EC9	U1EC8	vi
WITH TILDE	d	U0129	U0128	kl vi
WITH DOT BELOW	d	U1ECB	U1ECA	vi
WITH OGONEK	d	U012F	U012E	lt
WITH MACRON	d	U012B	U012A	lv [Cyr] [ar]
LATIN LETTER J	b	U006A	U004A	+
WITH ACUTE	d	–	–	[Cyr]
WITH CIRCUMFLEX	d	U0135	U0134	eo
WITH CARON	d	U01F0	–	[Cyr]
LATIN LETTER K	b	U006B	U004B	+
WITH ACUTE	d	U1E31	U1E30	[Cyr]
WITH GRAVE	d	–	–	[Cyr]

Name	Type	Position for lowercase/uppercase in ISO/IEC 10646-1		Languages used
WITH CIRCUMFLEX	d	–	–	[Cyr]
WITH CARON	d	U01E9	U01E8	[Cyr]
WITH DOT ABOVE	d	–	–	[he]
WITH DOT BELOW	d	U1E33	U1E32	[Cyr]
WITH COMMA BELOW [c]	d	–	–	lv [Cyr]
WITH MACRON	d	–	–	[Cyr]
LATIN LETTER K WITH HOOK	s	U0199	U0198	ha
LATIN SMALL LETTER KRA	s	U0138		kl
LATIN LETTER L	b	U006C	U004C	+
WITH ACUTE	d	U013A	U0139	[Cyr]
WITH CIRCUMFLEX	d	–	–	[Cyr]
WITH CIRCUMFLEX BELOW	d	U1E3D	U1E3C	ve
WITH CARON	d	U013E	U013D	
WITH DOT BELOW	d	U1E37	U1E36	
WITH COMMA BELOW [d]	d	–	–	lv [Cyr]
WITH STROKE	s	U0142	U0141	pl
LATIN LETTER M	b	U006D	U004D	+
WITH ACUTE	d	U1E3F	U1E3E	
WITH CIRCUMFLEX	d	–	–	lb
WITH DOT ABOVE	d	U1E41	U1E40	
WITH DOT BELOW	d	U1E43	U1E42	
LATIN LETTER N	b	U006E	U004E	+
WITH ACUTE	d	U0144	U0143	pl smi [Cyr]
WITH GRAVE	d	–	–	[Cyr]
WITH BREVE	d	–	–	[Cyr]
WITH CIRCUMFLEX	d	–	–	lb [Cyr]
WITH CIRCUMFLEX BELOW	d	U1E4B	U1E4A	hz ve
WITH CARON	d	U0148	U0147	cs sk
WITH TILDE	d	U00F1	U00D1	br es eu gl
WITH DOT ABOVE	d	U1E45	U1E44	ve [Cyr]
WITH DOT BELOW	d	U1E47	U1E46	[Cyr]
WITH COMMA BELOW [e]	d	–	–	lv [Cyr]
WITH MACRON	d	–	–	[Cyr]
PRECEDED BY APOSTROPHE	d	U0149	–	
FOLLOWED BY APOSTROPHE	d	–	–	ts
LATIN LETTER ENG	s	U014B	U014A	se
LATIN LETTER O	b	U006F	U004F	+
WITH ACUTE	d	U00F3	U00D3	af ca cs cy da es fo ga gd gl hu is it nl no pl pt qal sk sl ss vi [Cyr]
WITH GRAVE	d	U00F2	U00D2	ca cy fur gd it no pt qal rm vi [Cyr]
WITH BREVE	d	U014F	U014E	
WITH CIRCUMFLEX	d	U00F4	U00D4	af cy de fr fy kl no nso pt qal sk sl tn vi [Cyr]
WITH CIRCUMFLEX AND ACUTE	d	U1ED1	U1ED0	vi
WITH CIRCUMFLEX AND GRAVE	d	U1ED3	U1ED2	vi
WITH CIRCUMFLEX AND HOOK ABOVE	d	U1ED5	U1ED4	vi

Name	Type	Position for lowercase/uppercase in ISO/IEC 10646-1		Languages used
WITH CIRCUMFLEX AND TILDE	d	U1ED7	U1ED6	vi
WITH CIRCUMFLEX AND DOT BELOW	d	U1ED9	U1ED8	vi
WITH CARON	d	U01D2	U01D1	
WITH DIAERESIS	d	U00F6	U00D6	af cy de et fi fy hu is lb nl rm sv tr [Cyr]
WITH DIAERESIS AND DOT BELOW	d	–	–	[Cyr]
WITH DOUBLE ACUTE	d	U0151	U0150	hu [Cyr]
WITH HOOK ABOVE	d	U1ECF	U1ECE	vi
WITH TILDE	d	U00F5	U00D5	et pt vi
WITH DOT BELOW	d	U1ECD	U1ECC	vi
WITH OGONEK	d	U01EB	U01EA	
WITH MACRON	d	U014D	U014C	lv [Cyr]
WITH HORN	d	U01A1	U01A0	vi
WITH HORN AND ACUTE	d	U1EDB	U1EDA	vi
WITH HORN AND GRAVE	d	U1EDD	U1EDC	vi
WITH HORN AND HOOK ABOVE	d	U1EDF	U1EDE	vi
WITH HORN AND TILDE	d	U1EE1	U1EE0	vi
WITH HORN AND DOT BELOW	d	U1EE3	U1EE2	vi
WITH STROKE	s	U00F8	U00D8	da fo is kl no smi
WITH STROKE AND ACUTE	s d	U01FF	U01FE	
LATIN LIGATURE OE	s	U0153	U0152	fr [Cyr]
LATIN LETTER P	b	U0070	U0050	+
WITH ACUTE	d	U1E55	U1E54	[Cyr]
WITH GRAVE	d	–	–	[Cyr]
WITH DOT ABOVE	d	U1E57	U1E56	[he]
LATIN LETTER Q	b	U0071	U0051	+
LATIN LETTER R	b	U0072	U0052	+
WITH ACUTE	d	U0155	U0154	sk
WITH CARON	d	U0159	U0158	cs sk
WITH DOT ABOVE	d	U1E59	U1E58	
WITH DOT BELOW	d	U1E5B	U1E5A	
WITH CEDILLA	d	U0157	U0156	lv
LATIN LETTER S	b	U0073	U0053	+
WITH ACUTE	d	U015B	U015A	pl [he]
WITH GRAVE	d	–	–	[Cyr] [he]
WITH CIRCUMFLEX	d	U015D	U015C	eo [Cyr]
WITH CARON	d	U0161	U0160	cs et hr lt lv nso sk sl smi tn [Cyr] [ar] [he]
WITH DOT ABOVE	d	U1E61	U1E60	
WITH DOT BELOW	d	U1E63	U1E62	[ar] [he]
WITH CEDILLA	d	U015F	U015E	tr
WITH COMMA BELOW	d	–	–	mo ro [Cyr]
LATIN SMALL LETTER SHARP S	s	U00DF		de
LATIN LETTER T	b	U0074	U0054	+
WITH GRAVE	d	–	–	[Cyr]
WITH CIRCUMFLEX BELOW	d	U1E71	U1E70	hz ve
WITH CARON	d	U0165		(cs) (sk) [Cyr]

Name	Type	Position for lowercase/uppercase in ISO/IEC 10646-1		Languages used
WITH CARON	d		U0164	cs sk [Cyr]
WITH DIAERESIS	d	U1E97	–	[ar]
WITH DOT ABOVE	d	U1E6B	U1E6A	
WITH DOT BELOW	d	U1E6D	U1E6C	[ar] [he]
WITH COMMA BELOW [f]	d	–	–	mo ro [Cyr]
WITH LINE BELOW	d	U1E6F	U1E6E	[ar]
FOLLOWED BY APOSTROPHE	d	–		cs sk
WITH STROKE	s	U0167	U0166	se
LATIN LETTER U	b	U0075	U0055	+
WITH ACUTE	d	U00FA	U00DA	af ca cs cy da es fo fy ga gl hu is it kl nl pt qal sk vi [Cyr] [ar]
WITH GRAVE	d	U00F9	U00D9	br cy fr fur gd it qal vi [Cyr]
WITH BREVE	d	U016D	U016C	eo [Cyr]
WITH CIRCUMFLEX	d	U00FB	U00DB	af cy fr fy kl qal tr [Cyr]
WITH CARON	d	U01D4	U01D3	
WITH RING ABOVE	d	U016F	U016E	cs [Cyr]
WITH DIAERESIS	d	U00FC	U00DC	br ca cy de es et fr fy gl hu lb nl pt rm tr [Cyr]
WITH DIAERESIS AND DOT BELOW	d	–	–	[Cyr]
WITH DOUBLE ACUTE	d	U0171	U0170	hu [Cyr]
WITH HOOK ABOVE	d	U1EE7	U1EE6	vi
WITH TILDE	d	U0169	U0168	kl vi
WITH DOT ABOVE	d	–	–	[Cyr]
WITH DOT BELOW	d	U1EE5	U1EE4	vi
WITH OGONEK	d	U0173	U0172	lt
WITH MACRON	d	U016B	U016A	lt lv [Cyr] [ar]
WITH MACRON AND DOT BELOW	d	–	–	[Cyr]
WITH HORN	d	U01B0	U01AF	vi
WITH HORN AND ACUTE	d	U1EE9	U1EE8	vi
WITH HORN AND GRAVE	d	U1EEB	U1EEA	vi
WITH HORN AND HOOK ABOVE	d	U1EED	U1EEC	vi
WITH HORN AND TILDE	d	U1EEF	U1EEE	vi
WITH HORN AND DOT BELOW	d	U1EF1	U1EF0	vi
LATIN LETTER V	b	U0076	U0056	+
WITH TILDE	d	U1E7D	U1E7C	
WITH DOT BELOW	d	U1E7F	U1E7E	
LATIN LETTER W	b	U0077	U0057	+
WITH ACUTE	d	U1E83	U1E82	cy
WITH GRAVE	d	U1E81	U1E80	cy
WITH CIRCUMFLEX	d	U0175	U0174	cy [he]
WITH DIAERESIS	d	U1E85	U1E84	cy
WITH DOT ABOVE	d	U1E87	U1E86	[he]
WITH DOT BELOW	d	U1E89	U1E88	
LATIN LETTER X	b	U0078	U0058	+
WITH DIAERESIS	d	U1E8D	U1E8C	
WITH DOT ABOVE	d	U1E8B	U1E8A	
LATIN LETTER Y	b	U0079	U0059	+
WITH ACUTE	d	U00FD	U00DD	af cs cy is fo sk vi

Name	Type	Position for lowercase/uppercase in ISO/IEC 10646-1		Languages used
WITH GRAVE	d	U1EF3	U1EF2	cy vi [Cyr] [ar]
WITH CIRCUMFLEX	d	U0177	U0176	cy
WITH DIAERESIS	d	U00FF	U0178	cy fr [Cyr]
WITH HOOK ABOVE	d	U1EF7	U1EF6	vi
WITH TILDE	d	U1EF9	U1EF8	vi
WITH DOT ABOVE	d	U1E8F	U1E8E	
WITH DOT BELOW	d	U1EF5	U1EF4	vi
WITH MACRON	d	–	–	[Cyr]
LATIN LETTER Z	b	U007A	U005A	+
WITH ACUTE	d	U017A	U0179	pl [Cyr]
WITH BREVE	d	–	–	[Cyr]
WITH CIRCUMFLEX	d	U1E91	U1E90	[Cyr]
WITH CARON	d	U017E	U017D	cs hr lt lv sk sl smi [Cyr]
WITH CARON AND CEDILLA	d	–	–	[Cyr]
WITH DIAERESIS	d	–	–	[Cyr]
WITH DOT ABOVE	d	U017C	U017B	mt pl
WITH DOT BELOW	d	U1E93	U1E92	[ar]
WITH CEDILLA	d	–	–	[Cyr]
WITH MACRON	d	–	–	[Cyr]
LATIN LETTER THORN	b	U00FE	U00DE	is

[a] LATIN CAPITAL LETTER D WITH STROKE and LATIN CAPITAL LETTER ETH have identical form.
[b] The character is called LATIN LETTER G WITH CEDILLA in ISO/IEC 10646-1, positions U0123 and U0122.
[c] The character is called LATIN LETTER K WITH CEDILLA in ISO/IEC 10646-1, positions U0137 and U0136.
[d] The character is called LATIN LETTER L WITH CEDILLA in ISO/IEC 10646-1, positions U013C and U013B.
[e] The character is called LATIN LETTER N WITH CEDILLA in ISO/IEC 10646-1, positions U0146 and U0145.
[f] The character is called LATIN LETTER T WITH CEDILLA in ISO/IEC 10646-1, positions U0163 and U0162.

Annex E
(informative)

Languages using the Latin alphabet

Table E.1 lists languages using the Latin alphabet. The alpha-2 symbols are taken from ISO 639 ([8] in the bibliography); the alpha-3 symbols are taken from ISO 639-2 ([9] in the bibliography) (the terminology code is used when this differs from the bibliographic code). For some languages that are not included in ISO 639 or ISO 639-2, ad hoc symbols starting with "qa" have been assigned for reference.

Table E.2 lists alpha-2 and alpha-3 symbols and English names of languages in alphabetical order by alpha-2 and alpha-3 symbols.

Table E.1 — Alphabetical list by English names of language

alpha-2 code	alpha-3 code	English name of language	French name of language	alpha-2 code	alpha-3 code	English name of language	French name of language
af	afr	**Afrikaans**	afrikaans	ie	ile	**Interlingue**	interlingue
ak	aka	**Akan**	akan	iu	iku	**Inuktitut**	inuktitut
sq	sqi	**Albanian**	albanais	ik	ipk	**Inupiaq; Inupiak**	inupiaq
ay	aym	**Aymara**	aymara	ga	gle	**Irish; Irish Gaelic**	irlandais; gaélique d'Irlande; gaélique irlandais
az	aze	**Azerbaijani**	azéri; azerbaïdjanais				
eu	eus	**Basque**	basque	it	ita	**Italian**	italien
bi	bis	**Bislama**	bichlamar; bislama	jv	jaw	**Javanese**	javanais
br	bre	**Breton**	breton	kr	kau	**Kanuri**	kanouri
ca	cat	**Catalan**	catalan	ki	kik	**Kikuyu; Gikuyu**	kikuyu
ch	cha	**Chamorro**	chamorro	rw	kin	**Kinyarwanda; Rwanda**	kinyarwanda; rwanda
kw	cor	**Cornish**	cornique				
co	cos	**Corsican**	corse	rn	run	**Kirundi; Rundi**	kirundi; rundi
hr	hrv	**Croatian**	croate	kg	kon	**Kongo; Kikongo**	kongo
cs	ces	**Czech**	tchèque	ku	kur	**Kurdish**	kurde
da	dan	**Danish**	danois	kj	kua	**Kwanyama**	kwanyama
nl	nld	**Dutch**	néerlandais	–	qal	**Ladin**	ladin
–	efi	**Efik**	efik	–	lad	**Ladino**	ladino; judéo-espagnol
en	eng	**English**	anglais				
eo	epo	**Esperanto**	espéranto	la	lat	**Latin**	latin
et	est	**Estonian**	estonien	lv	lav	**Latvian; Lettish**	letton; lettonien
ee	ewe	**Ewe**	éwé	ln	lin	**Lingala**	lingala
fo	fao	**Faroese; Faeroese**	féroïen	lt	lit	**Lithuanian**	lituanien
fj	fij	**Fiji; Fijian**	fidji; fidjien	lu	lub	**Luba; Luba-Katanga**	tchiluba; luba-katanga
fi	fin	**Finnish**	finnois				
fr	fra	**French**	français	lg	lug	**Luganda; Ganda**	louganda; ganda
fy	fry	**Frisian**	frison	lb	ltz	**Luxemburgish**	luxembourgeois
–	fur	**Friulian**	frioulan	mg	mlg	**Malagasy**	malagache
ff	ful	**Fulah; Fula; Fulani; Fulfulde**	peol	ms	msa	**Malay**	malais
				mt	mlt	**Maltese**	maltais
gl	glg	**Galician; Gallegan**	galicien	–	man	**Mandingo**	mandingue
de	deu	**German**	allemand	gv	glv	**Manx Gaelic**	gaélique de l'Ile de Man; manx
–	gil	**Gilbertese**	kiribati				
kl	kal	**Greenlandic; Kalaallisut**	groenlandais	mi	mri	**Maori**	maori
ha	hau	**Hausa**	haoussa	mo	mol	**Moldavian**	moldave
–	haw	**Hawaiian**	hawaïen	na	nau	**Nauru; Nauruan**	nauruen; nauri
hz	her	**Herero**	herero	nd	nde	**Ndebele, North**	ndébéle du Nord
ho	hmo	**Hiri Motu**	hiri motu; hiri motou	nr	nbl	**Ndebele, South**	ndébéle du Sud
hu	hun	**Hungarian**	hongrois	ng	ndo	**Ndonga**	ndonga
is	isl	**Icelandic**	islandais	no	nor	**Norwegian**	norvégien
–	ibo	**Igbo**	igbo	oc	oci	**Occitan; Provençal**	occitan; provençal
id	ind	**Indonesian**	indonésien	oj	oji	**Ojibway**	ojibwa
ia	ina	**Interlingua**	interlingua	pl	pol	**Polish**	polonais

ISO 12199:2000(E)

Table E.1 (*continued*)

alpha-2 code	alpha-3 code	English name of language	French name of language
pt	por	**Portuguese**	portugais
qu	que	**Quechua**	quechua; quichua
rm	roh	**Raeto-Romance**	rhéto-roman
ro	ron	**Romanian**	roumain
–	smi	**Sami**	sami
se	sme	**Sami, Northern**	sami du Nord
sm	smo	**Samoan**	samoan
sg	sag	**Sango; Sangho**	sango
sc	srd	**Sardinian**	sarde
gd	gla	**Scottish Gaelic**	gaélique de l'Écosse; gaélique écossais
sh	–	**Serbo-Croatian**	serbo-croate
st	sot	**Sesotho**	sesotho; sotho du Sud
tn	tsn	**Setswana; Tswana**	setchwana; tswana
sn	sna	**Shona; chiShona**	shona
ss	ssw	**Siswati; Swati**	siswati; swati; swazi
sk	slk	**Slovak**	slovaque
sl	slv	**Slovenian**	slovène
so	som	**Somali**	somali
–	wen	**Sorbian**	sorabe
–	nso	**Sotho, Northern**	sotho du Nord
es	spa	**Spanish; Castilian**	espagnol, castillan

alpha-2 code	alpha-3 code	English name of language	French name of language
su	sun	**Sundanese**	soundanais
sw	swa	**Swahili; Kiswahili**	swahili
sv	swe	**Swedish**	suédois
tl	tgl	**Tagalog**	tagal; tagalog
ty	tah	**Tahitian**	tahitien
to	ton	**Tonga; Tongan**	tonga
ts	tso	**Tsonga**	tsonga
tr	tur	**Turkish**	turc
–	tvl	**Tuvalu**	tuvalu
tw	twi	**Twi; Tshi**	twi; tchi
ug	uig	**Uyghur**	ouigour
uz	uzb	**Uzbek**	ouzbek
ve	ven	**Venda**	venda
vi	vie	**Vietnamese**	vietnamien
vo	vol	**Volapük**	volapük
cy	cym	**Welsh**	gallois
wo	wol	**Wolof**	wolof
xh	xho	**Xhosa**	xhosa
yo	yor	**Yoruba**	yorouba; yoruba
za	zha	**Zhuang; Chwang; Chuang**	zhuang; jouang
zu	zul	**Zulu**	zoulou

Table E.2 — Alphabetical list by language symbols

alpha-2 and -3 code	English name of language
af afr	Afrikaans
ak aka	Akan
ay aym	Aymara
az aze	Azerbaijani
bi bis	Bislama
br bre	Breton
ca cat	Catalan
ces	Czech
ch cha	Chamorro
co	Corsican
cor	Cornish
cos	Corsican
cs	Czech
cy cym	Welsh
da dan	Danish
de deu	German
ee	Ewe
efi	Efik
en eng	English
eo epo	Esperanto
es	Spanish; Castilian
est et	Estonian
eu eus	Basque
ewe	Ewe
fao	Faroese; Faeroese
ff	Fulah; Fula; Fulani; Fulfulde
fi	Finnish
fij	Fiji; Fijian
fin	Finnish
fj	Fiji; Fijian
fo	Faroese; Faeroese
fr fra	French
fry	Frisian

alpha-2 and -3 code	English name of language
ful	Fulah; Fula; Fulani; Fulfulde
fur	Friulian
fy	Frisian
ga	Irish; Irish Gaelic
gd	Scottish Gaelic
gil	Gilbertese
gl	Galician; Gallegan
gla	Scottish Gaelic
gle	Irish; Irish Gaelic
glg	Galician; Gallegan
glv gv	Manx Gaelic
ha hau	Hausa
haw	Hawaiian
her	Herero
hmo ho	Hiri Motu
hr hrv	Croatian
hu hun	Hungarian
hz	Herero
ia	Interlingua
ibo	Igbo
id	Indonesian
ie	Interlingue
ik	Inupiaq; Inupiak
iku	Inuktitut
ile	Interlingue
ina	Interlingua
ind	Indonesian
ipk	Inupiaq; Inupiak
is isl	Icelandic
it ita	Italian
iu	Inuktitut
jaw jv	Javanese
kal	Greenlandic

Table E.2 (continued)

alpha-2 and -3 code	English name of language
kau	Kanuri
kg	Kongo; Kikongo
ki kik	Kikuyu; Gikuyu
kin	Kinyarwanda; Rwanda
kj	Kwanyama
kl	Greenlandic
kon	Kongo; Kikongo
kr	Kanuri
ku	Kurdish
kua	Kwanyama
kur	Kurdish
kw	Cornish
la	Latin
lad	Ladino
lat	Latin
lav	Latvian; Lettish
lb	Luxemburgish
lg	Luganda; Ganda
lin	Lingala
lit	Lithuanian
ln	Lingala
lt	Lithuanian
ltz	Luxemburgish
lu lub	Luba; Luba-Katanga
lug	Luganda; Ganda
lv	Latvian; Lettish
man	Mandingo
mg	Malagasy
mi	Maori
mlg	Malagasy
mlt	Maltese
mo mol	Moldavian
mri	Maori
ms msa	Malay
mt	Maltese
na nau	Nauru; Nauruan
nbl	Ndebele, South
nd nde	Ndebele, North
ndo ng	Ndonga
nl nld	Dutch
no nor	Norwegian
nr	Ndebele, South
nso	Sotho, Northern
oc oci	Occitan; Provençal
oj oji	Ojibway
pl pol	Polish
por pt	Portuguese
qal	Ladin
qu que	Quechua
rm	Raeto-Romance
rn	Kirundi; Rundi
ro	Romanian

alpha-2 and -3 code	English name of language
roh	Raeto-Romance
ron	Romanian
run	Kirundi; Rundi
rw	Kinyarwanda; Rwanda
sag	Sango; Sangho
sc	Sardinian
se	Sami, Northern
sg	Sango; Sangho
sh	Serbo-Croatian
sk	Slovak
sl	Slovenian
slk	Slovak
slv	Slovenian
sm	Samoan
sme	Sami, Northern
smi	Sami
smo	Samoan
sn sna	Shona; chiShona
so som	Somali
sot	Sesotho
spa	Spanish; Castilian
sq sqi	Albanian
srd	Sardinian
ss ssw	Siswati; Swati
st	Sesotho
su sun	Sundanese
sv	Swedish
sw swa	Swahili; Kiswahili
swe	Swedish
tah	Tahitian
tgl tl	Tagalog
tn	Setswana; Tswana
to ton	Tonga; Tongan
tr	Turkish
ts	Tsonga
tsn	Setswana; Tswana
tso	Tsonga
tur	Turkish
tvl	Tuvalu
tw twi	Twi; Tshi
ty	Tahitian
ug uig	Uyghur
uz uzb	Uzbek
ve ven	Venda
vi vie	Vietnamese
vo vol	Volapük
wen	Sorbian
wo wol	Wolof
xh xho	Xhosa
yo yor	Yoruba
za zha	Zhuang; Chwang; Chuang
zu zul	Zulu

ISO 12199:2000(E)

Annex F
(informative)

Alphabetical sequences and character repertoires

Most languages deviate from the rules that are laid down in this International Standard. Table F.1 formalizes rules given in national standards and other documents that will have to be followed in addition to, or instead of, the rules given in this International Standard to produce alphabetically ordered lists in these languages. Where no reference is given, the information is derived from dictionaries, etc.

Dictionaries and other alphabetically arranged lists in many languages may be based on different ordering rules and even different characters from those given below or those specified in national standards.

Deviations are grouped as follows:

①	deviations as to first-ordering-level values; see clause 5;
②	deviations as to other ordering rules; see clauses 6 to 8;
③	letters with diacritical marks and special letters that are native to the language; within parentheses, items that are used in loan words and names only; see 5.3 and 6.2;
④	basic Latin letters that are not native to the language (but may be used in, for example, names); within parentheses, items that are used in loan words; see 5.2, item b).

Lowercase letters are used to represent both lowercase and uppercase letters. See Table D.1 in annex D for the correspondence between lowercase and uppercase letters.

Basic Latin letters are shown in their typographical form. Most other letters are shown in their typographical form followed by the position in ISO/IEC 10646-1. Some letters are replaced by #1, #2, etc. followed by the position and character name in ISO/IEC 10646-1. U???? indicates that the character is not included in ISO/IEC 10646-1.

The following notation is used:

+	sequence of items following a letter;
=	items that are treated as equivalent;
→	deviations from the rules of character equivalence given in 5.3.

Table F.2 lists the character repertoires of some internationally standardized transliteration systems in Latin script.

NOTE 1　The languages that are included in Tables F.1 and F.2 are those for which information has been made available.

NOTE 2　In Table F.1 "Variants" are, for example, geographical, temporal or linguistic variants. For references, see the bibliography.

Table F.1 — Alphabetical sequence and character repertoires

Language	Variants; References	Deviations and other information on alphabet (① first level; ② other levels; ③ native letters; ④ non-native)					
Afrikaans af / afr		① –					
		② –					
		③ á U00E1	é U00E9	è U00E8	ê U00EA	ë U00EB	í U00ED
		î U00EE	ï U00EF	ó U00F3	ô U00F4	ö U00F6	ú U00FA
		û U00FB	ý U00FD	(à) U00E0	(ä) U00E4	(ç) U00E7	
		(ì) U00EC	(ò) U00F2	(ù) U00F9			
		④ þ					
Albanian sq / sqi		① c + ç　　d + dh　　e + ë　　g + gj　　l + ll　　n + nj　　r + rr						
		s + sh　　t + th　　x + xh　　z + zh						
		② –					
		③ ç U00E7　　ë U00EB					
		④ þ					

Language	Variants; References	Deviations and other information on alphabet (① first level; ② other levels; ③ native letters; ④ non-native)				
Basque eu / eus		① l + ll n + ñ t + tt ② – ③ ñ U00F1 (ç) U00E7 (ü) U00FC ④ þ				
Breton br / bre		① c + ch + c'h n + ñ ② – ③ â U00E2 ê U00EA ñ U00F1 ù U00F9 ü U00FC ④ þ (c) (q) (x)				
Catalan ca / cat		① l + ll l.l or l·l is ordered between lk and lm. ② – ③ á U00E1 à U00E0 ç U00E7 é U00E9 è U00E8 í U00ED ï U00EF ó U00F3 ò U00F2 ú U00FA ü U00FC ④ þ				
Cornish kw / cor		① – ② – ③ – ④ þ (x) (z)				
Corsican co / cos		① g + ghj z + chj ② – ③ à U00E0 è U00E8 ì U00EC ò U00F2 ù U00F9 ④ w x y þ j is used only in the polygraphs ghj and chj				
Croatian hr / hrv		① c + č + ć d + dž + đ l + lj n + nj s + š z + ž ② – ③ č U010D ć U0107 đ U0111 š U0161 ž U017E ④ þ (q) (w) (x) (y)				
Czech cs / ces		① c + č h + ch r + ř s + š z + ž ② – ③ á U00E1 č U010D ď U???? Ď U010E é U00E9 ě U011B í U00ED ň U0148 ó U00F3 ř U0159 š U0161 ť U???? Ť U0164 ú U00FA ů U016F ý U00FD ž U017E ď (lowercase) corresponds to Ď (uppercase) and is sometimes represented as U010F LATIN SMALL LETTER D WITH CARON. ť (lowercase) corresponds to Ť (uppercase) and is sometimes represented as U0165 LATIN SMALL LETTER T WITH CARON. ④ þ (w)				
Danish da / dan	DS 377 and Dansk sprognævn (Danish Language Council)	① z + æ + ø + å ü = y ű = y ä = æ ö = ø ö = ø þ → th aa → å (except when the two as belong to separate syllables) Word-by-word ordering is given priority. Numbers shall be transcribed to words before ordering. ② Uppercase letter before lowercase letter. ③ á U00E1 å U00E5 æ U00E6 é U00E9 í U00ED ó U00F3 ø U00F8 ú U00FA (ä) U00E4 (ö) U00F6 (ö) U0151 (ü) U00FC (ű) U0171 ④ þ				
Dutch nl / nld		① – The digraph ij may be treated as equivalent to y, especially in lists of names. ② – ③ á U00E1 à U00E0 ä U00E4 é U00E9 è U00E8 ê U00EA ë U00EB í U00ED ì U00EC ï U00EF ó U00F3 ò U00F2 ö U00F6 ú U00FA ù U00F9 ü U00FC The digraph ij/IJ was previously regarded as a ligature. The initial uppercase form is IJ (not "Ij"). ④ þ				

Language	Variants; References	Deviations and other information on alphabet (① first level; ② other levels; ③ native letters; ④ non-native)
English en / eng	BS 1749	① þ → th Letter-by-letter and word-by-word ordering have equal status. ② – ③ (à) U00E0 (æ) U00E6 (ç) U00E7 (é) U00E9 (è) U00E8 (ë) U00EB (ï) U00EF (ñ) U00F1 (ö) U00F6 (ô) U00F4 (œ) U0153 ④ þ
Esperanto eo / epo		① c + ĉ g + ĝ h + ĥ j + ĵ s + ŝ u + ŭ ② – ③ ĉ U0109 ĝ U011D ĥ U0125 ĵ U0135 ŝ U015D ŭ U016D ④ þ (q) (w) (x) (y)
Estonian et / est		① s + š (z) + õ + ä + ö + ü ② – ③ ä U00E4 ö U00F6 õ U00F5 š U0161 ü U00FC ④ z þ (c) (q) (w) (x) (y)
Faroese; Faeroese fo / fao		① a + á d + ð i + í o + ó u + ú y + ý z + æ + ø ② – ③ á U00E1 æ U00E6 ð U00F0 í U00ED ó U00F3 ø U00F8 ú U00FA ý U00FD ④ þ
Finnish fi / fin	SFS 4600	① š = s ž = z z + å + ä + ö w = v ü = y ø = ö þ → th ② – ③ ä U00E4 ö U00F6 (å) U00E5 (š) U0161 (ž) U017E ④ þ (b) (c) (f) (q) (w)
French fr / fra		① – ② Diacritics from right to left for each word. ③ à U00E0 â U00E2 æ U00E6 ç U00E7 é U00E9 è U00E8 ê U00EA ë U00EB î U00EE ï U00EF ô U00F4 œ U0153 ù U00F9 û U00FB ü U00FC ÿ U00FF ④ þ
Frisian fy / fry		① ij = y ② – ③ à U00E0 â U00E2 ä U00E4 é U00E9 ê U00EA ë U00EB ï U00EF ô U00F4 ö U00F6 ú U00FA û U00FB ü U00FC ④ þ
Friulian – / fur		① a + à + á + â e + è i + ì o + ò u + ù ② – ③ á U00E1 à U00E0 â U00E2 è U00E8 ì U00EC ò U00F2 ù U00F9 ④ þ
Galician; Gallegan gl / glg		① n + ñ ② – ③ á U00E1 é U00E9 í U00ED ñ U00F1 ó U00F3 ú U00FA ü U00FC ④ þ
German de / deu	DIN 5007 ON A 2725	① þ → th Word-by-word ordering specified. ② Ordering according to diacritical marks: "Umlaut" first. ③ à U00E0 â U00E2 ä U00E4 é U00E9 è U00E8 ê U00EA ë U00EB ï U00EF ô U00F4 ö U00F6 ß U00DF ü U00FC ④ þ

Language	Variants; References	Deviations and other information on alphabet (① first level; ② other levels; ③ native letters; ④ non-native)					
Green-landic; Kalaallisut kl / kal		① z + æ + ø + å κ = q					
		② –					
		③ á U00E1	â U00E2	å U00E5	ã U00E3	æ U00E6	é U00E9
		í U00ED	î U00EE	ī U0129	κ U0138	ô U00F4	ø U00F8
		ú U00FA	û U00FB	ū U0169			
		κ (U0138 LATIN SMALL LETTER KRA) has been replaced in the orthography by the letter q. The uppercase letter corresponding to LATIN SMALL LETTER KRA is K' (LATIN CAPITAL LETTER K FOLLOWED BY APOSTROPHE).					
		④ þ					
Hausa ha / hau		① b + #1 d + #2 k + #3					
		② –					
		③ #1 U0253 B WITH HOOK		#2 U0257 D WITH HOOK		#3 U0199 K WITH HOOK	
		④ j q v x þ					
Herero hz / her		① a + e + i + o + u +					
		h + k + m + n + #2 + p + r + s + t + #3 + v + w + y + z +					
		b + d + f + g + j + l +					
		c + q + x +					
		mb + nd + n#1 + ng + ny + ndj + tj					
		② –					
		③ #1 U1E13 D WITH CIRCUMFLEX BELOW		#2 U1E4B N WITH CIRCUMFLEX BELOW			
		#3 U1E71 T WITH CIRCUMFLEX BELOW					
		④ c q x þ					
Hungarian hu / hun		① c + cs	d + dz + dzs	g + gy	l + ly	n + ny	o + ö
		s + sz	t + ty	u + ü	z + zs	ő = ö	ű = ü
		② –					
		③ á U00E1	é U00E9	í U00ED	ó U00F3	ö U00F6	ő U0151
		ú U00FA	ü U00FC	ű U0171			
		④ þ (q) (w) (x)					
Icelandic is / isl	Íslensk mál-nefnd (Icelandic Language Council)	① a + á	d + ð	e + é	i + í	o + ó	u + ú	y + ý
		þ + æ + ö + (å) ä = æ ø = ö					
		② –					
		③ á U00E1	æ U00E6	ð U00F0	é U00E9	í U00ED	ó U00F3
		ö U00F6	ú U00FA	ý U00FD	(ä) U00E4	(å) U00E5	(ø) U00F8
		④ c q w z					
Irish; Irish Gaelic ga / gle		① –					
		② –					
		③ á U00E1	é U00E9	í U00ED	ó U00F3	ú U00FA	
		④ þ (k)					
		Irish also uses a Gaelic variant of the Latin script. This script has not been included in the character set of this International Standard.					
Italian it / ita		① –					
		② –					
		③ à U00E0	é U00E9	è U00E8	í U00ED	ì U00EC	î U00EE
		ï U00EF	ó U00F3	ò U00F2	ú U00FA	ù U00F9	
		④ þ					
Ladin – / qal		① –					
		② –					
		③ á U00E1	à U00E0	â U00E2	é U00E9	è U00E8	ê U00EA
		í U00ED	ì U00EC	î U00EE	ó U00F3	ò U00F2	ô U00F4
		ú U00FA	ù U00F9	û U00FB			
		④ þ					
Latin la / lat		① –					
		② –					
		③ –					
		④ j w þ (k)					

ISO 12199:2000(E)

Language	Variants; References	Deviations and other information on alphabet (① first level; ② other levels; ③ native letters; ④ non-native)					
Latvian; Lettish lv / lav		① c + č g + ġ k + ķ l + ļ n + ņ r + ŗ s + š z + ž ② — ③ ā U0101 č U010D ē U0113 ġ U???? G WITH COMMA ABOVE ī U012B ķ U???? K WITH COMMA BELOW ļ U???? L WITH COMMA BELOW ņ U???? N WITH COMMA BELOW ō U014D ŗ U???? R WITH COMMA BELOW š U0161 ū U016B ž U017E ġ (lowercase) corresponds to Ģ (uppercase) (G WITH COMMA ABOVE/BELOW). ④ þ (q) (w) (x) (y)					
Lithuanian lt / lit		① c + č s + š z + ž y = i ② Ogonek before dot above. ③ ą U0105 č U010D ę U0119 ė U0117 į U012F š U0161 ų U0173 ū U016B ž U017E ④ þ (q) (w) (x)					
Luxem-burgish lb / ltz		① — ② — ③ ä U00E4 é U00E9 ë U00EB #1 U???? M WITH CARON #2 U???? N WITH CARON ö U00F6 ü U00FC ④ þ					
Maltese mt / mlt		① f + ġ h + ħ n + għ y + ż ċ = c ② — ③ ċ U010B ġ U0121 ħ U0127 ż U017C (à) U00E0 (è) U00E8 (ì) U00EC (î) U00EE (ò) U00F2 (ù) U00F9 ④ þ (c) (y)					
Moldavian [a] mo / mol		① a + ă + â i + î s + ş t + ţ ② — ③ ă U0103 â U00E2 î U00EE ş U???? S WITH COMMA BELOW ţ U???? T WITH COMMA BELOW ④ þ					
Ndebele — / qan (nde, nbl)		① — ② — ③ — ④ h þ					
Norwegian no / nor	Bokmål and Nynorsk; NS 4103	① z + æ + ø + å ü = y ű = y ä = æ ö = ø ő = ø þ → th Digits after letters. ② — ③ à U00E0 å U00E5 æ U00E6 é U00E9 è U00E8 ê U00EA ó U00F3 ò U00F2 ô U00F4 ø U00F8 (ä) U00E4 (ö) U00F6 (ü) U00FC ④ þ					
Occitan; Provençal oc / oci		① — ② — ③ ç U00E7 ï U00EF ④ k w þ					
Polish pl / pol		① a + ą c + ć e + ę l + ł n + ń o + ó s + ś z + ź + ż ② — ③ ą U0105 ć U0107 ę U0119 ł U0142 ń U0144 ó U00F3 ś U015B ź U017A ż U017C ④ þ (q) (v) (x)					
Portuguese pt / por		① — ② — ③ á U00E1 à U00E0 â U00E2 ã U00E3 ç U00E7 é U00E9 è U00E8 ê U00EA í U00ED ó U00F3 ò U00F2 ô U00F4 õ U00F5 ú U00FA ü U00FC ④ þ					

Language	Variants; References	Deviations and other information on alphabet (① first level; ② other levels; ③ native letters; ④ non-native)
Raeto-Romance rm / roh		① – ② – ③ à U00E0 â U00E2 è U00E8 ê U00EA ò U00F2 ö U00F6 ü U00FC ④ þ
Romanian [a] ro / ron		① a + ă + â i + î s + ş t + ţ ② – ③ ă U0103 â U00E2 î U00EE ş U???? S WITH COMMA BELOW ţ U???? T WITH COMMA BELOW ④ þ
Sami – / smi	Northern Sami se / sme	① a + á c + č d + đ n + ŋ s + š t + t z + ž + (æ) + (ø) + (å) (ä) = (æ) (ö) = (ø) ② – ③ á U00E1 č U010D đ U0111 ŋ U014B š U0161 t U0167 ž U017E (ä) U00E4 (æ) U00E6 (ö) U00F6 (ø) U00F8 ④ þ (q) (w) (x) (y)
	Lule Sami	① a + á n + ń z + å + ä ② – ③ á U00E1 å U00E5 ä U00E4 ń U0144 ④ þ (c) (q) (w) (x) (y) (z)
	Inari Sami	① c + č d + đ s + š z + ž + ä á = ä ② – ③ á U00E1 â U00E2 ä U00E4 č U010D đ U0111 š U0161 ž U017E ④ þ (q) (w) (x)
	Southern Sami	① z + æ + ø + å (ä) = æ (ö) = ø ② – ③ å U00E5 æ U00E6 ø U00F8 (ä) U00E4 (ö) U00F6 ④ þ (c) (q) (w) (x) (z)
Scottish Gaelic gd / gla		① – ② – ③ á U00E1 à U00E0 é U00E9 è U00E8 í U00ED ì U00EC ó U00F3 ò U00F2 ù U00F9 ④ þ (j) (k) (q) (v) (w) (x) (y) (z)
Setswana; Tswana tn / tsn		① e + ê o + ô s + š ② – ③ ê U00EA ô U00F4 š U0161 ④ c q v x z þ
Siswati; Swati ss / ssw		① – ② – ③ á U00E1 é U00E9 ó U00F3 ④ q r x þ
Slovak sk / slk		① a + ä c + č h + ch r + ř s + š z + ž ② – ③ á U00E1 ä U00E4 č U010D ď U???? Ď U010E é U00E9 í U00ED ň U0148 ó U00F3 ô U00F4 ŕ U0155 ř U0159 š U0161 ť U???? Ť U0164 ú U00FA ý U00FD ž U017E ď (lowercase) corresponds to Ď (uppercase) and is sometimes represented as U010F LATIN SMALL LETTER D WITH CARON. ť (lowercase) corresponds to Ť (uppercase) and is sometimes represented as U0165 LATIN SMALL LETTER T WITH CARON. ④ þ (w)
Slovenian sl / slv		① c + č s + š z + ž ② – ③ č U010D é U00E9 ê U00EA ó U00F3 ô U00F4 š U0161 ž U017E ④ þ (q) (w) (x) (y)

Language	Variants; References	Deviations and other information on alphabet (① first level; ② other levels; ③ native letters; ④ non-native)
Somali so / som		① d + dh k + kh s + sh ② – ③ – ④ p v z þ
Sotho, Northern – / nso		① e + ê o + ô s + š ② – ③ ê U00EA ô U00F4 š U0161 ④ þ (c) (q) (v) (z)
Spanish; Castilian es / spa	Real Academia de la Lengua Española (new alphabetical sequence)	① n + ñ ② – ③ á U00E1 é U00E9 í U00ED ñ U00F1 ó U00F3 ú U00FA ü U00FC ④ þ (k) (w)
	Traditional alphabetical sequence	① c + ch l + ll n + ñ ② – ③ á U00E1 é U00E9 í U00ED ñ U00F1 ó U00F3 ú U00FA ü U00FC ④ þ (k) (w)
Swahili; Kiswahili sw / swa		① b + ch + d + dh g + gh n + ng' + ny s + sh + t + th ② – ③ – ④ þ
Swedish sv / swe	TN 34	① z + å + ä + ö v = w æ = a ü = y ø = ö ② – ③ ä U00E4 å U00E5 é U00E9 ö U00F6 (æ) U00E6 (ø) U00F8 (ü) U00FC ④ þ
Tsonga ts / tso		① n + n' ② – ③ n' U???? N FOLLOWED BY APOSTROPHE ④ þ

Turkish
 tr / tur
		① c + ç g + ğ h + ı + i o + ö s + ş u + ü
		 ② –
		 ③ â U00E2 ç U00E7 ğ U011F ı U0131 İ U0130 ö U00F6
		    ş U015F û U00FB ü U00FC
		    For details about ı/I and i/İ, see 5.2, item b), second paragraph after note 1.
		 ④ þ (q) (w) (x)

Venda
 ve / ven
		① b + #1 + d k + #2 + l m + #3 + n + #4 s + #5 + t
		 ② –
		 ③ #1 U1E13 D WITH CIRCUMFLEX BELOW #2 U1E3D L WITH CIRCUMFLEX BELOW
		    #3 U1E4B N WITH CIRCUMFLEX BELOW #4 U1E45 N WITH DOT ABOVE
		    #5 U1E71 T WITH CIRCUMFLEX BELOW
		 ④ c j q þ

Vietnamese
 vi / vie
		① a + ă + â c + ch d + đ e + ê g + gh + gi k + kh
		    n + ng + ngh + nh o + ô + #31 t + th + tr u + #40
		 ② Ordering according to diacritical marks indicating tone: grave, acute, hook above, tilde, dot below.
		 ③ à U00E0 á U00E1 #01 U1EA3 A WITH HOOK ABOVE ã U00E3
		    #02 U1EA1 A WITH DOT BELOW ă U0103
		    #03 U1EB1 A WITH BREVE AND GRAVE #04 U1EAF A WITH BREVE AND ACUTE
		    #05 U1EB3 A WITH BREVE AND HOOK ABOVE
		    #06 U1EB5 A WITH BREVE AND TILDE #07 U1EB7 A WITH BREVE AND TILDE
		    â U00E2 #08 U1EA7 A WITH CIRCUMFLEX AND GRAVE

Language	Variants; References	Deviations and other information on alphabet (① first level; ② other levels; ③ native letters; ④ non-native)						
Vietnamese (continued)		(continued) #09 U1EA5 A WITH CIRCUMFLEX AND ACUTE #10 U1EA9 A WITH CIRCUMFLEX AND HOOK ABOVE #11 U1EAB A WITH CIRCUMFLEX AND TILDE #12 U1EAD A WITH CIRCUMFLEX AND DOT BELOW đ U0111 è U00E8 é U00E9 #13 U1EBB E WITH HOOK ABOVE #14 U1EBD E WITH TILDE #15 U1EB9 E WITH DOT BELOW ê U00EA #16 U1EC1 E WITH CIRCUMFLEX AND GRAVE #17 U1EBF E WITH CIRCUMFLEX AND ACUTE #18 U1EC3 E WITH CIRCUMFLEX AND HOOK ABOVE #19 U1EC5 E WITH CIRCUMFLEX AND TILDE #20 U1EC7 E WITH CIRCUMFLEX AND DOT BELOW ì U00EC í U00ED #21 U1EC9 I WITH HOOK ABOVE #22 U0129 I WITH TILDE #23 U1ECB I WITH DOT BELOW ò U00F2 ó U00F3 #24 U1ECF O WITH HOOK ABOVE õ U00F5 #25 U1ECD O WITH DOT BELOW ô U00F4 #26 U1ED3 O WITH CIRCUMFLEX AND GRAVE #27 U1ED1 O WITH CIRCUMFLEX AND ACUTE #28 1ED5 O WITH CIRCUMFLEX AND HOOK ABOVE #29 U1ED7 O WITH CIRCUMFLEX AND TILDE #30 U1ED9 O WITH CIRCUMFLEX AND DOT BELOW #31 U01A1 O WITH HORN #32 U1EDD O WITH HORN AND GRAVE #33 U1EDB O WITH HORN AND ACUTE #34 U1EDF O WITH HORN AND HOOK ABOVE #35 U1EE1 O WITH HORN AND TILDE #36 U1EE3 O WITH HORN AND DOT BELOW ù U00F9 ú U00FA #37 U1EE7 U WITH HOOK ABOVE #38 U0169 U WITH TILDE #39 U1EE5 U WITH DOT BELOW #40 U01B0 U WITH HORN #41 U1EEB U WITH HORN AND GRAVE #42 U1EE9 U WITH HORN AND ACUTE #43 U1EED U WITH HORN AND HOOK ABOVE #44 U1EEF U WITH HORN AND TILDE #45 U1EF1 U WITH HORN AND DOT BELOW #46 U1EF3 Y WITH GRAVE ý U00FD #47 U1EF7 Y WITH HOOK ABOVE #48 U1EF9 Y WITH TILDE #49 U1EF5 Y WITH DOT BELOW ④ f j v z þ						
Welsh cy / cym		① c + ch d + dd f + ff g + ng l + ll p + ph r + rh t + th ② — ③ á U00E1 à U00E0 â U00E2 ä U00E4 é U00E9 è U00E8 ê U00EA ë U00EB í U00ED ì U00EC î U00EE ï U00EF ó U00F3 ò U00F2 ô U00F4 ö U00F6 ú U00FA ù U00F9 û U00FB ü U00FC #1 U1E83 W WITH ACUTE #2 U1E81 W WITH GRAVE ŵ U0175 #3 U1E85 W WITH DIAERESIS ý U00FD #4 U1EF3 Y WITH GRAVE #5 U0177 Y WITH CIRCUMFLEX ÿ U00FF ④ þ (q) (v) (x)						
Xhosa xh / xho		① — ② — ③ — ④ þ						
Zulu zu / zul		① — ② — ③ — ④ þ (r)						
^a The Moldavian language and the Romanian language are identical.								

Table F.2 — Character repertoires of some transliteration standards

Language, Script	References	Character repertoire
Greek [el]	ISO 843	*Table 1:* a v g d e z ī U012B (or ī⁻) th i k l m n x o p r s t y f ch ps ō U014D (or o⁻) *Table 2:* a ai ái aï av af ay v g ng gk nx nch d e ei éi eï ev ef ey z i iv if iy th i k l m b mp n nt x o oi ói oï ou oy p r s t y yi f ch ps o
Cyrillic [Cyr] a	ISO 9	*Table 1:* a b v g d đ U0111 #1 U01F5 G WITH ACUTE e ë U00EB ê U00EA ž U017E z #2 U1E91 Z WITH CIRCUMFLEX i ì U00EC ï U00EF #3 U01F0 J WITH CARON j k l #4 U???? L WITH CIRCUMFLEX m n #5 U???? N WITH CIRCUMFLEX o p r s t ć U0107 #6 U1E31 K WITH ACUTE u ŭ U016D f h c č U010D #7 U???? D WITH CIRCUMFLEX š U0161 ŝ U015D " U0022 y ' U0027 è U00E8 û U00FB â U00E2 ' U02BC *Table 2:* #1 U???? G WITH GRAVE ě U011B #2 U01CE A WITH CARON #3 U???? F WITH GRAVE #4 U1EF3 Y WITH GRAVE *Table 3:* a ä U00E4 #01 U???? A WITH DIAERESIS AND DOT BELOW ă U0103 ā U0101 æ U00E6 á U00E1 å U00E5 b v g #02 U01F5 G WITH ACUTE ġ U0121 ğ U011F #03 U1E25 H WITH DOT BELOW d e ĕ U0115 ë U00EB ž U017E #04 U???? Z WITH CARON AND CEDILLA #05 U???? Z WITH MACRON #06 U???? Z WITH BREVE z #07 U???? Z WITH CEDILLA #08 U???? Z WITH DIAERESIS ź U017A i ī U012B í U00ED î U00EE j ì U00EC ï U00EF #09 U01D0 I WITH CARON #10 U01F0 J WITH CARON #11 U???? J WITH ACUTE k . #12 U???? K WITH DOT BELOW #13 U???? K WITH CIRCUMFLEX #14 U???? K WITH CARON #15 U???? K WITH MACRON #16 U???? K WITH COMMA BELOW #17 U???? K WITH GRAVE q l #18 U013A L WITH ACUTE #19 U???? L WITH COMMA BELOW m n #20 U???? N WITH COMMA BELOW #21 U1E47 N WITH DOT BELOW #22 U1E45 N WITH DOT ABOVE #23 U???? N WITH GRAVE ń U0144 #24 U???? N WITH BREVE #25 U???? N WITH MACRON o ö U00F6 ô U00F4 ő U0151 #26 U???? O WITH DIAERESIS AND DOT BELOW ò U00F2 ó U00F3 ō U014D œ U0153 p #27 U1E55 P WITH ACUTE #28 U???? P WITH GRAVE r s #29 U???? S WITH COMMA BELOW #30 U???? S WITH GRAVE t #31 U???? T WITH GRAVE #32 U0165 T WITH CARON #33 U???? T WITH COMMA BELOW u ü U00FC ū U016B ŭ U016D ű U0171 ú U00FA #34 U???? U WITH DIAERESIS AND DOT BELOW ù U00F9 #35 U???? U WITH DOT ABOVE #36 U???? U WITH MACRON AND DOT BELOW w f h #37 U1E29 H WITH CEDILLA c #38 U???? X WITH MACRON #39 U???? D WITH CIRCUMFLEX č U010D ç U00E7 #40 U???? C WITH DOT BELOW #41 U???? C WITH DIAERESIS ĉ U0109 #42 U???? C WITH GRAVE #43 U???? C WITH BREVE #44 U???? C WITH BREVE AND COMMA BELOW š U0161 ŝ U015D " U0022 y ÿ U00FF #45 U???? Y WITH MACRON ' U0027 è U00E8 #46 U???? A WITH DOUBLE ACUTE à U00E0 û U00FB ů U016F â U00E2

Language, Script	References	Character repertoire
Arabic [ar]	ISO 233	*Table 1:* ' U02BC 'â U02BC+00E2 b t #01 U1E6F T WITH LINE BELOW #02 U01E7 G WITH CARON #03 U1E25 H WITH DOT BELOW #04 U1E96 H WITH LINE BELOW d #05 U1E0F D WITH LINE BELOW r z s š U0161 #06 U1E63 S WITH DOT BELOW #07 U1E0D D WITH DOT BELOW #08 U1E6D T WITH DOT·BELOW #09 U1E93 Z WITH DOT BELOW ' U02BB ġ U0121 f q k l m n h #10 U1E97 T WITH DIAERESIS w y #11 U1EF3 Y WITH GRAVE *Table 2:* a a' á' U00E1+02BC á U00E1 ā U0101 u ú U00FA i í U00ED
	ISO 233-2	*Table 1:* ' U02BC b t #01 U1E6F T WITH LINE BELOW #02 U01E7 G WITH CARON #03 U1E25 H WITH DOT BELOW #04 U1E96 H WITH LINE BELOW d #05 U1E0F D WITH LINE BELOW r z s š U0161 #06 U1E63 S WITH DOT BELOW #07 U1E0D D WITH DOT BELOW #08 U1E6D T WITH DOT BELOW #09 U1E93 Z WITH DOT BELOW ' U02BB ġ U0121 f q k l m h #10 U1E97 T WITH DIAERESIS w y *Table 2:* a ā U0101 â U00E2 á U00E1 u ū U016B i ī U012B
Hebrew [he]	ISO 259-2	*Table 1:* ' U02BC b #1 U1E03 B WITH DOT ABOVE g d h #2 U1E23 H WITH DOT ABOVE w z #3 U1E25 H WITH DOT BELOW #4 U1E6D T WITH DOT BELOW y #5 U???? K WITH DOT ABOVE l m n s ' U01BB p #6 U1E57 P WITH DOT ABOVE #7 U1E63 S WITH DOT BELOW q r #8 U???? S WITH GRAVE š U0161 ś U015B t *Table 2:* a e i o u ŵ U0175 #1 U1E87 W WITH DOT ABOVE
^a The symbol Cyr is used for the Cyrillic alphabet.		

Annex G
(normative)

Formal description of the rules of the main body of this International Standard

```
%% ISO 12199 Table
section ISO12199CollatingSymbols
%% This table is based on
   ISO14651_1999_TABLE1
%%
%% Digits
<0>
<1>
<2>
<3>
<4>
<5>
<6>
<7>
<8>
<9>
%% Latin letters
<a>
<b>
<c>
<d>
<e>
<f>
<g>
<h>
<i>
<j>
<k>
<l>
<m>
<n>
<o>
<p>
<q>
<r>
<s>
<t>
<u>
<v>
<w>
<x>
<y>
<z>
<thorn>
%% Collating symbols for the second ordering
   level key
<BLANK>
<ACUTE>
<GRAVE>
<BREVE>
<CIRCUMFLEX>
<CIRCUMFLEX_BELOW>
<CARON>
<RING_ABOVE>
<DIAERESIS>
<DOUBLE_ACUTE>
<HOOK_ABOVE>
<TILDE>
<DOT_ABOVE>
<DOT_BELOW>
<CEDILLA>
<COMMA>
<OGONEK>
<MACRON>
<MACRON_BELOW>
<HORN>
<MODIFIED1>
<MODIFIED2>
<MODIFIED3>
<SUBSTITUTE>
%% Collating symbols for the third ordering
   level key
<LOWERCASE>
```

```
<UPPERCASE>
%%
section ISO12199Digits
<U0030> <0>; <BLANK>; <LOWERCASE>; <U0030> %
   DIGIT ZERO
<U2070> <0>; <BLANK>; <LOWERCASE>; <U2070> %
   SUPERSCRIPT ZERO
<U2080> <0>; <BLANK>; <LOWERCASE>; <U2080> %
   SUBSCRIPT ZERO
<U0031> <1>; <BLANK>; <LOWERCASE>; <U0031> %
   DIGIT ONE
<U00B9> <1>; <BLANK>; <LOWERCASE>; <U00B9> %
   SUPERSCRIPT ONE
<U2081> <1>; <BLANK>; <LOWERCASE>; <U2081> %
   SUBSCRIPT ONE
<U0032> <2>; <BLANK>; <LOWERCASE>; <U0032> %
   DIGIT TWO
<U00B2> <2>; <BLANK>; <LOWERCASE>; <U00B2> %
   SUPERSCRIPT TWO
<U2082> <2>; <BLANK>; <LOWERCASE>; <U2082> %
   SUBSCRIPT TWO
<U0033> <3>; <BLANK>; <LOWERCASE>; <U0033> %
   DIGIT THREE
<U00B3> <3>; <BLANK>; <LOWERCASE>; <U00B3> %
   SUPERSCRIPT THREE
<U2083> <3>; <BLANK>; <LOWERCASE>; <U2083> %
   SUBSCRIPT THREE
<U0034> <4>; <BLANK>; <LOWERCASE>; <U0034> %
   DIGIT FOUR
<U2074> <4>; <BLANK>; <LOWERCASE>; <U2074> %
   SUPERSCRIPT FOUR
<U2084> <4>; <BLANK>; <LOWERCASE>; <U2084> %
   SUBSCRIPT FOUR
<U0035> <5>; <BLANK>; <LOWERCASE>; <U0035> %
   DIGIT FIVE
<U2075> <5>; <BLANK>; <LOWERCASE>; <U2075> %
   SUPERSCRIPT FIVE
<U2085> <5>; <BLANK>; <LOWERCASE>; <U2085> %
   SUBSCRIPT FIVE
<U0036> <6>; <BLANK>; <LOWERCASE>; <U0036> %
   DIGIT SIX
<U2076> <6>; <BLANK>; <LOWERCASE>; <U2076> %
   SUPERSCRIPT SIX
<U2086> <6>; <BLANK>; <LOWERCASE>; <U2086> %
   SUBSCRIPT SIX
<U0037> <7>; <BLANK>; <LOWERCASE>; <U0037> %
   DIGIT SEVEN
<U2077> <7>; <BLANK>; <LOWERCASE>; <U2077> %
   SUPERSCRIPT SEVEN
<U2087> <7>; <BLANK>; <LOWERCASE>; <U2087> %
   SUBSCRIPT SEVEN
<U0038> <8>; <BLANK>; <LOWERCASE>; <U0038> %
   DIGIT EIGHT
<U2078> <8>; <BLANK>; <LOWERCASE>; <U2078> %
   SUPERSCRIPT EIGHT
<U2088> <8>; <BLANK>; <LOWERCASE>; <U2088> %
   SUBSCRIPT EIGHT
<U0039> <9>; <BLANK>; <LOWERCASE>; <U0039> %
   DIGIT NINE
<U2079> <9>; <BLANK>; <LOWERCASE>; <U2079> %
   SUPERSCRIPT NINE
<U2089> <9>; <BLANK>; <LOWERCASE>; <U2089> %
   SUBSCRIPT NINE
%%
section ISO12199LatinLetters
<U0061> <a>; <BLANK>; <LOWERCASE>; <U0061> %
   LATIN SMALL LETTER A
<U0041> <a>; <BLANK>; <UPPERCASE>; <U0041> %
   LATIN CAPITAL LETTER A
<U00E1> <a>; <ACUTE>; <LOWERCASE>; <U00E1> %
   LATIN SMALL LETTER A WITH ACUTE
<U00C1> <a>; <ACUTE>; <UPPERCASE>; <U00C1> %
   LATIN CAPITAL LETTER A WITH ACUTE
```

```
<U00E0> <a>; <GRAVE>; <LOWERCASE>; <U00E0> %
   LATIN SMALL LETTER A WITH GRAVE
<U00C0> <a>; <GRAVE>; <UPPERCASE>; <U00C0> %
   LATIN CAPITAL LETTER A WITH GRAVE
<U0103> <a>; <BREVE>; <LOWERCASE>; <U0103> %
   LATIN SMALL LETTER A WITH BREVE
<U0102> <a>; <BREVE>; <UPPERCASE>; <U0102> %
   LATIN CAPITAL LETTER A WITH BREVE
<U1EAF> <a>; <BREVE><ACUTE>; <LOWERCASE>;
   <U1EAF> % LATIN SMALL LETTER A WITH BREVE
   AND ACUTE
<U1EAE> <a>; <BREVE><ACUTE>; <UPPERCASE>;
   <U1EAE> % LATIN CAPITAL LETTER A WITH
   BREVE AND ACUTE
<U1EB1> <a>; <BREVE><GRAVE>; <LOWERCASE>;
   <U1EB1> % LATIN SMALL LETTER A WITH BREVE
   AND GRAVE
<U1EB0> <a>; <BREVE><GRAVE>; <UPPERCASE>;
   <U1EB0> % LATIN CAPITAL LETTER A WITH
   BREVE AND GRAVE
<U1EB3> <a>; <BREVE><HOOK_ABOVE>;
   <LOWERCASE>; <U1EB3> % LATIN SMALL LETTER
   A WITH BREVE AND HOOK ABOVE
<U1EB2> <a>; <BREVE><HOOK_ABOVE>;
   <UPPERCASE>; <U1EB2> % LATIN CAPITAL
   LETTER A WITH BREVE AND HOOK ABOVE
<U1EB5> <a>; <BREVE><TILDE>; <LOWERCASE>;
   <U1EB5> % LATIN SMALL LETTER A WITH BREVE
   AND TILDE
<U1EB4> <a>; <BREVE><TILDE>; <UPPERCASE>;
   <U1EB4> % LATIN CAPITAL LETTER A WITH
   BREVE AND TILDE
<U1EB7> <a>; <BREVE><DOT_BELOW>; <LOWERCASE>;
   <U1EB7> % LATIN SMALL LETTER A WITH BREVE
   AND DOT BELOW
<U1EB6> <a>; <BREVE><DOT_BELOW>; <UPPERCASE>;
   <U1EB6> % LATIN CAPITAL LETTER A WITH
   BREVE AND DOT BELOW
<U00E2> <a>; <CIRCUMFLEX>; <LOWERCASE>;
   <U00E2> % LATIN SMALL LETTER A WITH
   CIRCUMFLEX
<U00C2> <a>; <CIRCUMFLEX>; <UPPERCASE>;
   <U00C2> % LATIN CAPITAL LETTER A WITH
   CIRCUMFLEX
<U1EA5> <a>; <CIRCUMFLEX><ACUTE>;
   <LOWERCASE>; <U1EA5> % LATIN SMALL LETTER
   A WITH CIRCUMFLEX AND ACUTE
<U1EA4> <a>; <CIRCUMFLEX><ACUTE>;
   <UPPERCASE>; <U1EA4> % LATIN CAPITAL
   LETTER A WITH CIRCUMFLEX AND ACUTE
<U1EA7> <a>; <CIRCUMFLEX><GRAVE>;
   <LOWERCASE>; <U1EA7> % LATIN SMALL LETTER
   A WITH CIRCUMFLEX AND GRAVE
<U1EA6> <a>; <CIRCUMFLEX><GRAVE>;
   <UPPERCASE>; <U1EA6> % LATIN CAPITAL
   LETTER A WITH CIRCUMFLEX AND GRAVE
<U1EA9> <a>; <CIRCUMFLEX><HOOK_ABOVE>;
   <LOWERCASE>; <U1EA9> % LATIN SMALL LETTER
   A WITH CIRCUMFLEX AND HOOK ABOVE
<U1EA8> <a>; <CIRCUMFLEX><HOOK_ABOVE>;
   <UPPERCASE>; <U1EA8> % LATIN CAPITAL
   LETTER A WITH CIRCUMFLEX AND HOOK ABOVE
<U1EAB> <a>; <CIRCUMFLEX><TILDE>;
   <LOWERCASE>; <U1EAB> % LATIN SMALL LETTER
   A WITH CIRCUMFLEX AND TILDE
<U1EAA> <a>; <CIRCUMFLEX><TILDE>;
   <UPPERCASE>; <U1EAA> % LATIN CAPITAL
   LETTER A WITH CIRCUMFLEX AND TILDE
<U1EAD> <a>; <CIRCUMFLEX><DOT_BELOW>;
   <LOWERCASE>; <U1EAD> % LATIN SMALL LETTER
   A WITH CIRCUMFLEX AND DOT BELOW
```

<U1EAC> <a>; <CIRCUMFLEX><DOT_BELOW>;
<UPPERCASE>; <U1EAC> % LATIN CAPITAL
LETTER A WITH CIRCUMFLEX AND DOT BELOW
<U01CE> <a>; <CARON>; <LOWERCASE>; <U01CE> %
LATIN SMALL LETTER A WITH CARON
<U01CD> <a>; <CARON>; <UPPERCASE>; <U01CD> %
LATIN CAPITAL LETTER A WITH CARON
<U00E5> <a>; <RING_ABOVE>; <LOWERCASE>;
<U00E5> % LATIN SMALL LETTER A WITH RING
ABOVE
<U00C5> <a>; <RING_ABOVE>; <UPPERCASE>;
<U00C5> % LATIN CAPITAL LETTER A WITH
RING ABOVE
<U01FB> <a>; <RING_ABOVE><ACUTE>;
<LOWERCASE>; <U01FB> % LATIN SMALL LETTER
A WITH RING ABOVE AND ACUTE
<U01FA> <a>; <RING_ABOVE><ACUTE>;
<UPPERCASE>; <U01FA> % LATIN CAPITAL
LETTER A WITH RING ABOVE AND ACUTE
<U00E4> <a>; <DIAERESIS>; <LOWERCASE>;
<U00E4> % LATIN SMALL LETTER A WITH
DIAERESIS
<U00C4> <a>; <DIAERESIS>; <UPPERCASE>;
<U00C4> % LATIN CAPITAL LETTER A WITH
DIAERESIS
<U01DF> <a>; <DIAERESIS><MACRON>;
<LOWERCASE>; <U01DF> % LATIN SMALL LETTER
A WITH DIAERESIS AND MACRON
<U01DE> <a>; <DIAERESIS><MACRON>;
<UPPERCASE>; <U01DE> % LATIN CAPITAL
LETTER A WITH DIAERESIS AND MACRON
<U1EA3> <a>; <HOOK_ABOVE>; <LOWERCASE>;
<U1EA3> % LATIN SMALL LETTER A WITH HOOK
ABOVE
<U1EA2> <a>; <HOOK_ABOVE>; <UPPERCASE>;
<U1EA2> % LATIN CAPITAL LETTER A WITH
HOOK ABOVE
<U00E3> <a>; <TILDE>; <LOWERCASE>; <U00E3> %
LATIN SMALL LETTER A WITH TILDE
<U00C3> <a>; <TILDE>; <UPPERCASE>; <U00C3> %
LATIN CAPITAL LETTER A WITH TILDE
<U1EA1> <a>; <DOT_BELOW>; <LOWERCASE>;
<U1EA1> % LATIN SMALL LETTER A WITH DOT
BELOW
<U1EA0> <a>; <DOT_BELOW>; <UPPERCASE>;
<U1EA0> % LATIN CAPITAL LETTER A WITH DOT
BELOW
<U0105> <a>; <OGONEK>; <LOWERCASE>; <U0105> %
LATIN SMALL LETTER A WITH OGONEK
<U0104> <a>; <OGONEK>; <UPPERCASE>; <U0104> %
LATIN CAPITAL LETTER A WITH OGONEK
<U0101> <a>; <MACRON>; <LOWERCASE>; <U0101> %
LATIN SMALL LETTER A WITH MACRON
<U0100> <a>; <MACRON>; <UPPERCASE>; <U0100> %
LATIN CAPITAL LETTER A WITH MACRON
<U00E6> "<a><e>"; <MODIFIED1><BLANK>;
<LOWERCASE><LOWERCASE>; <U0061><U0065> %
LATIN SMALL LETTER AE
<U00C6> "<a><e>"; <MODIFIED1><BLANK>;
<UPPERCASE><UPPERCASE>; <U0041><U0045> %
LATIN CAPITAL LETTER AE
<U01FD> "<a><e>"; <MODIFIED1><ACUTE>;
<LOWERCASE><LOWERCASE>; <U0061><U0065> %
LATIN SMALL LETTER AE WITH ACUTE
<U01FC> "<a><e>"; <MODIFIED1><ACUTE>;
<UPPERCASE><UPPERCASE>; <U0041><U0045> %
LATIN CAPITAL LETTER AE WITH ACUTE
<U01E3> "<a><e>"; <MODIFIED1><MACRON>;
<LOWERCASE><LOWERCASE>; <U0061><U0065> %
LATIN SMALL LETTER AE WITH MACRON
<U01E2> "<a><e>"; <MODIFIED1><MACRON>;
<UPPERCASE><UPPERCASE>; <U0041><U0045> %
LATIN CAPITAL LETTER AE WITH MACRON
<U0062> <b>; <BLANK>; <LOWERCASE>; <U0062> %
LATIN SMALL LETTER B
<U0042> <b>; <BLANK>; <UPPERCASE>; <U0042> %
LATIN CAPITAL LETTER B
<U1E03> <b>; <DOT_ABOVE>; <LOWERCASE>;
<U1E03> % LATIN SMALL LETTER B WITH DOT
ABOVE
<U1E02> <b>; <DOT_ABOVE>; <UPPERCASE>;
<U1E02> % LATIN CAPITAL LETTER B WITH DOT
ABOVE
<U1E05> <b>; <DOT_BELOW>; <LOWERCASE>;
<U1E05> % LATIN SMALL LETTER B WITH DOT
BELOW

<U1E04> <b>; <DOT_BELOW>; <UPPERCASE>;
<U1E04> % LATIN CAPITAL LETTER B WITH DOT
BELOW
<U0253> <b>; <MODIFIED1>; <LOWERCASE>;
<U0253> % LATIN SMALL LETTER B WITH HOOK
<U0181> <b>; <MODIFIED1>; <UPPERCASE>;
<U0181> % LATIN CAPITAL LETTER B WITH
HOOK
<U0063> <c>; <BLANK>; <LOWERCASE>; <U0063> %
LATIN SMALL LETTER C
<U0043> <c>; <BLANK>; <UPPERCASE>; <U0043> %
LATIN CAPITAL LETTER C
<U0107> <c>; <ACUTE>; <LOWERCASE>; <U0107> %
LATIN SMALL LETTER C WITH ACUTE
<U0106> <c>; <ACUTE>; <UPPERCASE>; <U0106> %
LATIN CAPITAL LETTER C WITH ACUTE
<U0109> <c>; <CIRCUMFLEX>; <LOWERCASE>;
<U0109> % LATIN SMALL LETTER C WITH
CIRCUMFLEX
<U0108> <c>; <CIRCUMFLEX>; <UPPERCASE>;
<U0108> % LATIN CAPITAL LETTER C WITH
CIRCUMFLEX
<U010D> <c>; <CARON>; <LOWERCASE>; <U010D> %
LATIN SMALL LETTER C WITH CARON
<U010C> <c>; <CARON>; <UPPERCASE>; <U010C> %
LATIN CAPITAL LETTER C WITH CARON
<U010B> <c>; <DOT_ABOVE>; <LOWERCASE>;
<U010B> % LATIN SMALL LETTER C WITH DOT
ABOVE
<U010A> <c>; <DOT_ABOVE>; <UPPERCASE>;
<U010A> % LATIN CAPITAL LETTER C WITH DOT
ABOVE
<U00E7> <c>; <CEDILLA>; <LOWERCASE>; <U00E7>
% LATIN SMALL LETTER C WITH CEDILLA
<U00C7> <c>; <CEDILLA>; <UPPERCASE>; <U00C7>
% LATIN CAPITAL LETTER C WITH CEDILLA
<U0188> <c>; <MODIFIED1>; <LOWERCASE>;
<U0188> % LATIN SMALL LETTER C WITH HOOK
<U0187> <c>; <MODIFIED1>; <UPPERCASE>;
<U0187> % LATIN CAPITAL LETTER C WITH
HOOK
<U0064> <d>; <BLANK>; <LOWERCASE>; <U0064> %
LATIN SMALL LETTER D
<U0044> <d>; <BLANK>; <UPPERCASE>; <U0044> %
LATIN CAPITAL LETTER D
<U1E13> <d>; <CIRCUMFLEX_BELOW>; <LOWERCASE>;
<U1E13> % LATIN SMALL LETTER D WITH
CIRCUMFLEX BELOW
<U1E12> <d>; <CIRCUMFLEX_BELOW>; <UPPERCASE>;
<U1E12> % LATIN CAPITAL LETTER D WITH
CIRCUMFLEX BELOW
<U010F> <d>; <CARON>; <LOWERCASE>; <U010F> %
LATIN SMALL LETTER D WITH CARON
<U010E> <d>; <CARON>; <UPPERCASE>; <U010E> %
LATIN CAPITAL LETTER D WITH CARON
<U1E0B> <d>; <DOT_ABOVE>; <LOWERCASE>;
<U1E0B> % LATIN SMALL LETTER D WITH DOT
ABOVE
<U1E0A> <d>; <DOT_ABOVE>; <UPPERCASE>;
<U1E0A> % LATIN CAPITAL LETTER D WITH DOT
ABOVE
<U1E0D> <d>; <DOT_BELOW>; <LOWERCASE>;
<U1E0D> % LATIN SMALL LETTER D WITH DOT
BELOW
<U1E0C> <d>; <DOT_BELOW>; <UPPERCASE>;
<U1E0C> % LATIN CAPITAL LETTER D WITH DOT
BELOW
<U1E0F> <d>; <MACRON_BELOW>; <LOWERCASE>;
<U1E0F> % LATIN SMALL LETTER D WITH LINE
BELOW
<U1E0E> <d>; <MACRON_BELOW>; <UPPERCASE>;
<U1E0E> % LATIN CAPITAL LETTER D WITH
LINE BELOW
<U0111> <d>; <MODIFIED1>; <LOWERCASE>;
<U0111> % LATIN SMALL LETTER D WITH
STROKE
<U0110> <d>; <MODIFIED1>; <UPPERCASE>;
<U0110> % LATIN CAPITAL LETTER D WITH
STROKE
<U0257> <d>; <MODIFIED2>; <LOWERCASE>;
<U0257> % LATIN SMALL LETTER D WITH HOOK
<U018A> <d>; <MODIFIED2>; <UPPERCASE>;
<U018A> % LATIN CAPITAL LETTER D WITH
HOOK
<U00F0> <d>; <MODIFIED3>; <LOWERCASE>;
<U00F0> % LATIN SMALL LETTER ETH

<U00D0> <d>; <MODIFIED3>; <UPPERCASE>;
<U00D0> % LATIN CAPITAL LETTER ETH
<U02A3> "<d><z>"; <BLANK><BLANK>;
<LOWERCASE><LOWERCASE>; <U0064><U007A> %
LATIN SMALL LETTER DZ DIGRAPH
<U01F3> "<d><z>"; <BLANK><BLANK>;
<LOWERCASE><LOWERCASE>; <U0064><U007A> %
LATIN SMALL LETTER DZ
<U01C6> "<d><z>"; <BLANK><CARON>;
<LOWERCASE><LOWERCASE>; <U0064><U017E> %
LATIN SMALL LETTER DZ WITH CARON
<U01F2> "<d><z>"; <BLANK><BLANK>;
<UPPERCASE><LOWERCASE>; <U0044><U007A> %
LATIN CAPITAL LETTER D WITH SMALL LETTER
Z
<U01F1> "<d><z>"; <BLANK><BLANK>;
<UPPERCASE><UPPERCASE>; <U0044><U005A> %
LATIN CAPITAL LETTER DZ
<U01C5> "<d><z>"; <BLANK><CARON>;
<UPPERCASE><LOWERCASE>; <U0044><U017E> %
LATIN CAPITAL LETTER D WITH SMALL LETTER
Z WITH CARON
<U01C4> "<d><z>"; <BLANK><CARON>;
<UPPERCASE><UPPERCASE>; <U0044><U017D> %
LATIN CAPITAL LETTER DZ WITH CARON
<U0065> <e>; <BLANK>; <LOWERCASE>; <U0065> %
LATIN SMALL LETTER E
<U0045> <e>; <BLANK>; <UPPERCASE>; <U0045> %
LATIN CAPITAL LETTER E
<U00E9> <e>; <ACUTE>; <LOWERCASE>; <U00E9> %
LATIN SMALL LETTER E WITH ACUTE
<U00C9> <e>; <ACUTE>; <UPPERCASE>; <U00C9> %
LATIN CAPITAL LETTER E WITH ACUTE
<U00E8> <e>; <GRAVE>; <LOWERCASE>; <U00E8> %
LATIN SMALL LETTER E WITH GRAVE
<U00C8> <e>; <GRAVE>; <UPPERCASE>; <U00C8> %
LATIN CAPITAL LETTER E WITH GRAVE
<U0115> <e>; <BREVE>; <LOWERCASE>; <U0115> %
LATIN SMALL LETTER E WITH BREVE
<U0114> <e>; <BREVE>; <UPPERCASE>; <U0114> %
LATIN CAPITAL LETTER E WITH BREVE
<U00EA> <e>; <CIRCUMFLEX>; <LOWERCASE>;
<U00EA> % LATIN SMALL LETTER E WITH
CIRCUMFLEX
<U00CA> <e>; <CIRCUMFLEX>; <UPPERCASE>;
<U00CA> % LATIN CAPITAL LETTER E WITH
CIRCUMFLEX
<U1EBF> <e>; <CIRCUMFLEX><ACUTE>;
<LOWERCASE>; <U1EBF> % LATIN SMALL LETTER
E WITH CIRCUMFLEX AND ACUTE
<U1EBE> <e>; <CIRCUMFLEX><ACUTE>;
<UPPERCASE>; <U1EBE> % LATIN CAPITAL
LETTER E WITH CIRCUMFLEX AND ACUTE
<U1EC1> <e>; <CIRCUMFLEX><GRAVE>;
<LOWERCASE>; <U1EC1> % LATIN SMALL LETTER
E WITH CIRCUMFLEX AND GRAVE
<U1EC0> <e>; <CIRCUMFLEX><GRAVE>;
<UPPERCASE>; <U1EC0> % LATIN CAPITAL
LETTER E WITH CIRCUMFLEX AND GRAVE
<U1EC3> <e>; <CIRCUMFLEX><HOOK_ABOVE>;
<LOWERCASE>; <U1EC3> % LATIN SMALL LETTER
E WITH CIRCUMFLEX AND HOOK ABOVE
<U1EC2> <e>; <CIRCUMFLEX><HOOK_ABOVE>;
<UPPERCASE>; <U1EC2> % LATIN CAPITAL
LETTER E WITH CIRCUMFLEX AND HOOK ABOVE
<U1EC5> <e>; <CIRCUMFLEX><TILDE>;
<LOWERCASE>; <U1EC5> % LATIN SMALL LETTER
E WITH CIRCUMFLEX AND TILDE
<U1EC4> <e>; <CIRCUMFLEX><TILDE>;
<UPPERCASE>; <U1EC4> % LATIN CAPITAL
LETTER E WITH CIRCUMFLEX AND TILDE
<U1EC7> <e>; <CIRCUMFLEX><DOT_BELOW>;
<LOWERCASE>; <U1EC7> % LATIN SMALL LETTER
E WITH CIRCUMFLEX AND DOT BELOW
<U1EC6> <e>; <CIRCUMFLEX><DOT_BELOW>;
<UPPERCASE>; <U1EC6> % LATIN CAPITAL
LETTER E WITH CIRCUMFLEX AND DOT BELOW
<U1E19> <e>; <CIRCUMFLEX_BELOW>; <LOWERCASE>;
<U1E19> % LATIN SMALL LETTER E WITH
CIRCUMFLEX BELOW
<U1E18> <e>; <CIRCUMFLEX_BELOW>; <UPPERCASE>;
<U1E18> % LATIN CAPITAL LETTER E WITH
CIRCUMFLEX BELOW
<U011B> <e>; <CARON>; <LOWERCASE>; <U011B> %
LATIN SMALL LETTER E WITH CARON

33

ISO 12199:2000(E)

<U011A> <e>; <CARON>; <UPPERCASE>; <U011A> %
LATIN CAPITAL LETTER E WITH CARON
<U00EB> <e>; <DIAERESIS>; <LOWERCASE>;
<U00EB> % LATIN SMALL LETTER E WITH
DIAERESIS
<U00CB> <e>; <DIAERESIS>; <UPPERCASE>;
<U00CB> % LATIN CAPITAL LETTER E WITH
DIAERESIS
<U1EBB> <e>; <HOOK_ABOVE>; <LOWERCASE>;
<U1EBB> % LATIN SMALL LETTER E WITH HOOK
ABOVE
<U1EBA> <e>; <HOOK_ABOVE>; <UPPERCASE>;
<U1EBA> % LATIN CAPITAL LETTER E WITH
HOOK ABOVE
<U1EBD> <e>; <TILDE>; <LOWERCASE>; <U1EBD> %
LATIN SMALL LETTER E WITH TILDE
<U1EBC> <e>; <TILDE>; <UPPERCASE>; <U1EBC> %
LATIN CAPITAL LETTER E WITH TILDE
<U0117> <e>; <DOT_ABOVE>; <LOWERCASE>;
<U0117> % LATIN SMALL LETTER E WITH DOT
ABOVE
<U0116> <e>; <DOT_ABOVE>; <UPPERCASE>;
<U0116> % LATIN CAPITAL LETTER E WITH DOT
ABOVE
<U1EB9> <e>; <DOT_BELOW>; <LOWERCASE>;
<U1EB9> % LATIN SMALL LETTER E WITH DOT
BELOW
<U1EB8> <e>; <DOT_BELOW>; <UPPERCASE>;
<U1EB8> % LATIN CAPITAL LETTER E WITH DOT
BELOW
<U0119> <e>; <OGONEK>; <LOWERCASE>; <U0119> %
LATIN SMALL LETTER E WITH OGONEK
<U0118> <e>; <OGONEK>; <UPPERCASE>; <U0118> %
LATIN CAPITAL LETTER E WITH OGONEK
<U0113> <e>; <MACRON>; <LOWERCASE>; <U0113> %
LATIN SMALL LETTER E WITH MACRON
<U0112> <e>; <MACRON>; <UPPERCASE>; <U0112> %
LATIN CAPITAL LETTER E WITH MACRON
<U0066> <f>; <BLANK>; <LOWERCASE>; <U0066> %
LATIN SMALL LETTER F
<U0046> <f>; <BLANK>; <UPPERCASE>; <U0046> %
LATIN CAPITAL LETTER F
<UFB00> "<f><f>"; <BLANK><BLANK>;
<LOWERCASE><LOWERCASE>; <U0066><U0066> %
LATIN SMALL LIGATURE FF
<UFB03> "<f><f><i>"; <BLANK><BLANK><BLANK>;
<LOWERCASE><LOWERCASE><LOWERCASE>;
<U0066><U0066><U0069> % LATIN SMALL
LIGATURE FFI
<UFB04> "<f><f><l>"; <BLANK><BLANK><BLANK>;
<LOWERCASE><LOWERCASE><LOWERCASE>;
<U0066><U0066><U006C> % LATIN SMALL
LIGATURE FFL
<UFB01> "<f><i>"; <BLANK><BLANK>;
<LOWERCASE><LOWERCASE>; <U0066><U0069> %
LATIN SMALL LIGATURE FI
<UFB02> "<f><l>"; <BLANK><BLANK>;
<LOWERCASE><LOWERCASE>; <U0066><U006C> %
LATIN SMALL LIGATURE FL
<U0067> <g>; <BLANK>; <LOWERCASE>; <U0067> %
LATIN SMALL LETTER G
<U0047> <g>; <BLANK>; <UPPERCASE>; <U0047> %
LATIN CAPITAL LETTER G
<U01F5> <g>; <ACUTE>; <LOWERCASE>; <U01F5> %
LATIN SMALL LETTER G WITH ACUTE
<U01F4> <g>; <ACUTE>; <UPPERCASE>; <U01F4> %
LATIN CAPITAL LETTER G WITH ACUTE
<U011F> <g>; <BREVE>; <LOWERCASE>; <U011F> %
LATIN SMALL LETTER G WITH BREVE
<U011E> <g>; <BREVE>; <UPPERCASE>; <U011E> %
LATIN CAPITAL LETTER G WITH BREVE
<U011D> <g>; <CIRCUMFLEX>; <LOWERCASE>;
<U011D> % LATIN SMALL LETTER G WITH
CIRCUMFLEX
<U011C> <g>; <CIRCUMFLEX>; <UPPERCASE>;
<U011C> % LATIN CAPITAL LETTER G WITH
CIRCUMFLEX
<U01E7> <g>; <CARON>; <LOWERCASE>; <U01E7> %
LATIN SMALL LETTER G WITH CARON
<U01E6> <g>; <CARON>; <UPPERCASE>; <U01E6> %
LATIN CAPITAL LETTER G WITH CARON
<U0121> <g>; <DOT_ABOVE>; <LOWERCASE>;
<U0121> % LATIN SMALL LETTER G WITH DOT
ABOVE

<U0120> <g>; <DOT_ABOVE>; <UPPERCASE>;
<U0120> % LATIN CAPITAL LETTER G WITH DOT
ABOVE
<U0123> <g>; <CEDILLA>; <LOWERCASE>; <U0123>
% LATIN SMALL LETTER G WITH CEDILLA
<U0122> <g>; <CEDILLA>; <UPPERCASE>; <U0122>
% LATIN CAPITAL LETTER G WITH CEDILLA
<U0260> <g>; <MODIFIED1>; <LOWERCASE>;
<U0260> % LATIN SMALL LETTER G WITH HOOK
<U0193> <g>; <MODIFIED1>; <UPPERCASE>;
<U0193> % LATIN CAPITAL LETTER G WITH
HOOK
<U0068> <h>; <BLANK>; <LOWERCASE>; <U0068> %
LATIN SMALL LETTER H
<U0048> <h>; <BLANK>; <UPPERCASE>; <U0048> %
LATIN CAPITAL LETTER H
<U0125> <h>; <CIRCUMFLEX>; <LOWERCASE>;
<U0125> % LATIN SMALL LETTER H WITH
CIRCUMFLEX
<U0124> <h>; <CIRCUMFLEX>; <UPPERCASE>;
<U0124> % LATIN CAPITAL LETTER H WITH
CIRCUMFLEX
<U1E23> <h>; <DOT_ABOVE>; <LOWERCASE>;
<U1E23> % LATIN SMALL LETTER H WITH DOT
ABOVE
<U1E22> <h>; <DOT_ABOVE>; <UPPERCASE>;
<U1E22> % LATIN CAPITAL LETTER H WITH DOT
ABOVE
<U1E25> <h>; <DOT_BELOW>; <LOWERCASE>;
<U1E25> % LATIN SMALL LETTER H WITH DOT
BELOW
<U1E24> <h>; <DOT_BELOW>; <UPPERCASE>;
<U1E24> % LATIN CAPITAL LETTER H WITH DOT
BELOW
<U1E29> <h>; <CEDILLA>; <LOWERCASE>; <U1E29>
% LATIN SMALL LETTER H WITH CEDILLA
<U1E28> <h>; <CEDILLA>; <UPPERCASE>; <U1E28>
% LATIN CAPITAL LETTER H WITH CEDILLA
<U1E96> <h>; <MACRON_BELOW>; <LOWERCASE>;
<U1E96> % LATIN SMALL LETTER H WITH LINE
BELOW
<U0127> <h>; <MODIFIED1>; <LOWERCASE>;
<U0127> % LATIN SMALL LETTER H WITH
STROKE
<U0126> <h>; <MODIFIED1>; <UPPERCASE>;
<U0126> % LATIN CAPITAL LETTER H WITH
STROKE
<U0069> <i>; <BLANK>; <LOWERCASE>; <U0069> %
LATIN SMALL LETTER I
<U0131> <i>; <BLANK>; <LOWERCASE>; <U0131> %
LATIN SMALL LETTER DOTLESS I
<U0049> <i>; <BLANK>; <UPPERCASE>; <U0049> %
LATIN CAPITAL LETTER I
<U0130> <i>; <BLANK>; <UPPERCASE>; <U0130> %
LATIN CAPITAL LETTER I WITH DOT ABOVE
<U00ED> <i>; <ACUTE>; <LOWERCASE>; <U00ED> %
LATIN SMALL LETTER I WITH ACUTE
<U00CD> <i>; <ACUTE>; <UPPERCASE>; <U00CD> %
LATIN CAPITAL LETTER I WITH ACUTE
<U00EC> <i>; <GRAVE>; <LOWERCASE>; <U00EC> %
LATIN SMALL LETTER I WITH GRAVE
<U00CC> <i>; <GRAVE>; <UPPERCASE>; <U00CC> %
LATIN CAPITAL LETTER I WITH GRAVE
<U012D> <i>; <BREVE>; <LOWERCASE>; <U012D> %
LATIN SMALL LETTER I WITH BREVE
<U012C> <i>; <BREVE>; <UPPERCASE>; <U012C> %
LATIN CAPITAL LETTER I WITH BREVE
<U00EE> <i>; <CIRCUMFLEX>; <LOWERCASE>;
<U00EE> % LATIN SMALL LETTER I WITH
CIRCUMFLEX
<U00CE> <i>; <CIRCUMFLEX>; <UPPERCASE>;
<U00CE> % LATIN CAPITAL LETTER I WITH
CIRCUMFLEX
<U01D0> <i>; <CARON>; <LOWERCASE>; <U01D0> %
LATIN SMALL LETTER I WITH CARON
<U01CF> <i>; <CARON>; <UPPERCASE>; <U01CF> %
LATIN CAPITAL LETTER I WITH CARON
<U00EF> <i>; <DIAERESIS>; <LOWERCASE>;
<U00EF> % LATIN SMALL LETTER I WITH
DIAERESIS
<U00CF> <i>; <DIAERESIS>; <UPPERCASE>;
<U00CF> % LATIN CAPITAL LETTER I WITH
DIAERESIS
<U1EC9> <i>; <HOOK_ABOVE>; <LOWERCASE>;
<U1EC9> % LATIN SMALL LETTER I WITH HOOK
ABOVE

<U1EC8> <i>; <HOOK_ABOVE>; <UPPERCASE>;
<U1EC8> % LATIN CAPITAL LETTER I WITH
HOOK ABOVE
<U0129> <i>; <TILDE>; <LOWERCASE>; <U0129> %
LATIN SMALL LETTER I WITH TILDE
<U0128> <i>; <TILDE>; <UPPERCASE>; <U0128> %
LATIN CAPITAL LETTER I WITH TILDE
<U1ECB> <i>; <DOT_BELOW>; <LOWERCASE>;
<U1ECB> % LATIN SMALL LETTER I WITH DOT
BELOW
<U1ECA> <i>; <DOT_BELOW>; <UPPERCASE>;
<U1ECA> % LATIN CAPITAL LETTER I WITH DOT
BELOW
<U012F> <i>; <OGONEK>; <LOWERCASE>; <U012F> %
LATIN SMALL LETTER I WITH OGONEK
<U012E> <i>; <OGONEK>; <UPPERCASE>; <U012E> %
LATIN CAPITAL LETTER I WITH OGONEK
<U012B> <i>; <MACRON>; <LOWERCASE>; <U012B> %
LATIN SMALL LETTER I WITH MACRON
<U012A> <i>; <MACRON>; <UPPERCASE>; <U012A> %
LATIN CAPITAL LETTER I WITH MACRON
<U0133> "<i><j>"; <BLANK><BLANK>;
<LOWERCASE><LOWERCASE>; <U0069><U006A> %
LATIN SMALL LIGATURE IJ
<U006A> <j>; <BLANK>; <LOWERCASE>; <U006A> %
LATIN SMALL LETTER J
<U004A> <j>; <BLANK>; <UPPERCASE>; <U004A> %
LATIN CAPITAL LETTER J
<U0135> <j>; <CIRCUMFLEX>; <LOWERCASE>;
<U0135> % LATIN SMALL LETTER J WITH
CIRCUMFLEX
<U0134> <j>; <CIRCUMFLEX>; <UPPERCASE>;
<U0134> % LATIN CAPITAL LETTER J WITH
CIRCUMFLEX
<U01F0> <j>; <CARON>; <LOWERCASE>; <U01F0> %
LATIN SMALL LETTER J WITH CARON
<U006B> <k>; <BLANK>; <LOWERCASE>; <U006B> %
LATIN SMALL LETTER K
<U004B> <k>; <BLANK>; <UPPERCASE>; <U004B> %
LATIN CAPITAL LETTER K
<U1E31> <k>; <ACUTE>; <LOWERCASE>; <U1E31> %
LATIN SMALL LETTER K WITH ACUTE
<U1E30> <k>; <ACUTE>; <UPPERCASE>; <U1E30> %
LATIN CAPITAL LETTER K WITH ACUTE
<U01E9> <k>; <CARON>; <LOWERCASE>; <U01E9> %
LATIN SMALL LETTER K WITH CARON
<U01E8> <k>; <CARON>; <UPPERCASE>; <U01E8> %
LATIN CAPITAL LETTER K WITH CARON
<U1E33> <k>; <DOT_BELOW>; <LOWERCASE>;
<U1E33> % LATIN SMALL LETTER K WITH DOT
BELOW
<U1E32> <k>; <DOT_BELOW>; <UPPERCASE>;
<U1E32> % LATIN CAPITAL LETTER K WITH DOT
BELOW
<U0137> <k>; <CEDILLA>; <LOWERCASE>; <U0137>
% LATIN SMALL LETTER K WITH CEDILLA
<U0136> <k>; <CEDILLA>; <UPPERCASE>; <U0136>
% LATIN CAPITAL LETTER K WITH CEDILLA
<U0199> <k>; <MODIFIED1>; <LOWERCASE>;
<U0199> % LATIN SMALL LETTER K WITH HOOK
<U0198> <k>; <MODIFIED1>; <UPPERCASE>;
<U0198> % LATIN CAPITAL LETTER K WITH
HOOK
<U0138> <k>; <MODIFIED2>; <LOWERCASE>;
<U0138> % LATIN SMALL LETTER KRA
<U006C> <l>; <BLANK>; <LOWERCASE>; <U006C> %
LATIN SMALL LETTER L
<U004C> <l>; <BLANK>; <UPPERCASE>; <U004C> %
LATIN CAPITAL LETTER L
<U013A> <l>; <ACUTE>; <LOWERCASE>; <U013A> %
LATIN SMALL LETTER L WITH ACUTE
<U0139> <l>; <ACUTE>; <UPPERCASE>; <U0139> %
LATIN CAPITAL LETTER L WITH ACUTE
<U1E3D> <l>; <CIRCUMFLEX_BELOW>; <LOWERCASE>;
<U1E3D> % LATIN SMALL LETTER L WITH
CIRCUMFLEX BELOW
<U1E3C> <l>; <CIRCUMFLEX_BELOW>; <UPPERCASE>;
<U1E3C> % LATIN CAPITAL LETTER L WITH
CIRCUMFLEX BELOW
<U013E> <l>; <CARON>; <LOWERCASE>; <U013E> %
LATIN SMALL LETTER L WITH CARON
<U013D> <l>; <CARON>; <UPPERCASE>; <U013D> %
LATIN CAPITAL LETTER L WITH CARON
<U1E37> <l>; <DOT_BELOW>; <LOWERCASE>;
<U1E37> % LATIN SMALL LETTER L WITH DOT
BELOW

34

```
<U1E36> <1>; <DOT_BELOW>; <UPPERCASE>;
    <U1E36> % LATIN CAPITAL LETTER L WITH DOT
    BELOW
<U013C> <1>; <CEDILLA>; <LOWERCASE>; <U013C>
    % LATIN SMALL LETTER L WITH CEDILLA
<U013B> <1>; <CEDILLA>; <UPPERCASE>; <U013B>
    % LATIN CAPITAL LETTER L WITH CEDILLA
<U0142> <1>; <MODIFIED1>; <LOWERCASE>;
    <U0142> % LATIN SMALL LETTER L WITH
    STROKE
<U0141> <1>; <MODIFIED1>; <UPPERCASE>;
    <U0141> % LATIN CAPITAL LETTER L WITH
    STROKE
<U01C9> "<1><j>"; <BLANK><BLANK>;
    <LOWERCASE><LOWERCASE>; <U006C><U006A> %
    LATIN SMALL LETTER LJ
<U01C8> "<1><j>"; <BLANK><BLANK>;
    <UPPERCASE><LOWERCASE>; <U004C><U006A> %
    LATIN CAPITAL LETTER L WITH SMALL LETTER
    J
<U01C7> "<1><j>"; <BLANK><BLANK>;
    <UPPERCASE><UPPERCASE>; <U004C><U004A> %
    LATIN CAPITAL LETTER LJ
<U006D> <m>; <BLANK>; <LOWERCASE>; <U006D> %
    LATIN SMALL LETTER M
<U004D> <m>; <BLANK>; <UPPERCASE>; <U004D> %
    LATIN CAPITAL LETTER M
<U1E3F> <m>; <ACUTE>; <LOWERCASE>; <U1E3F> %
    LATIN SMALL LETTER M WITH ACUTE
<U1E3E> <m>; <ACUTE>; <UPPERCASE>; <U1E3E> %
    LATIN CAPITAL LETTER M WITH ACUTE
<U1E41> <m>; <DOT_ABOVE>; <LOWERCASE>;
    <U1E41> % LATIN SMALL LETTER M WITH DOT
    ABOVE
<U1E40> <m>; <DOT_ABOVE>; <UPPERCASE>;
    <U1E40> % LATIN CAPITAL LETTER M WITH DOT
    ABOVE
<U1E43> <m>; <DOT_BELOW>; <LOWERCASE>;
    <U1E43> % LATIN SMALL LETTER M WITH DOT
    BELOW
<U1E42> <m>; <DOT_BELOW>; <UPPERCASE>;
    <U1E42> % LATIN CAPITAL LETTER M WITH DOT
    BELOW
<U006E> <n>; <BLANK>; <LOWERCASE>; <U006E> %
    LATIN SMALL LETTER N
<U004E> <n>; <BLANK>; <UPPERCASE>; <U004E> %
    LATIN CAPITAL LETTER N
<U0144> <n>; <ACUTE>; <LOWERCASE>; <U0144> %
    LATIN SMALL LETTER N WITH ACUTE
<U0143> <n>; <ACUTE>; <UPPERCASE>; <U0143> %
    LATIN CAPITAL LETTER N WITH ACUTE
<U1E4B> <n>; <CIRCUMFLEX_BELOW>; <LOWERCASE>;
    <U1E4B> % LATIN SMALL LETTER N WITH
    CIRCUMFLEX BELOW
<U1E4A> <n>; <CIRCUMFLEX_BELOW>; <UPPERCASE>;
    <U1E4A> % LATIN CAPITAL LETTER N WITH
    CIRCUMFLEX BELOW
<U0148> <n>; <CARON>; <LOWERCASE>; <U0148> %
    LATIN SMALL LETTER N WITH CARON
<U0147> <n>; <CARON>; <UPPERCASE>; <U0147> %
    LATIN CAPITAL LETTER N WITH CARON
<U00F1> <n>; <TILDE>; <LOWERCASE>; <U00F1> %
    LATIN SMALL LETTER N WITH TILDE
<U00D1> <n>; <TILDE>; <UPPERCASE>; <U00D1> %
    LATIN CAPITAL LETTER N WITH TILDE
<U1E45> <n>; <DOT_ABOVE>; <LOWERCASE>;
    <U1E45> % LATIN SMALL LETTER N WITH DOT
    ABOVE
<U1E44> <n>; <DOT_ABOVE>; <UPPERCASE>;
    <U1E44> % LATIN CAPITAL LETTER N WITH DOT
    ABOVE
<U1E47> <n>; <DOT_BELOW>; <LOWERCASE>;
    <U1E47> % LATIN SMALL LETTER N WITH DOT
    BELOW
<U1E46> <n>; <DOT_BELOW>; <UPPERCASE>;
    <U1E46> % LATIN CAPITAL LETTER N WITH DOT
    BELOW
<U0146> <n>; <CEDILLA>; <LOWERCASE>; <U0146>
    % LATIN SMALL LETTER N WITH CEDILLA
<U0145> <n>; <CEDILLA>; <UPPERCASE>; <U0145>
    % LATIN CAPITAL LETTER N WITH CEDILLA
<U014B> <n>; <MODIFIED1>; <LOWERCASE>;
    <U014B> % LATIN SMALL LETTER ENG
<U014A> <n>; <MODIFIED1>; <UPPERCASE>;
    <U014A> % LATIN CAPITAL LETTER ENG

<U01CC> "<n><j>"; <BLANK><BLANK>;
    <LOWERCASE><LOWERCASE>; <U006E><U006A> %
    LATIN SMALL LETTER NJ
<U01CB> "<n><j>"; <BLANK><BLANK>;
    <UPPERCASE><LOWERCASE>; <U004E><U006A> %
    LATIN CAPITAL LETTER N WITH SMALL LETTER
    J
<U01CA> "<n><j>"; <BLANK><BLANK>;
    <UPPERCASE><UPPERCASE>; <U004E><U004A> %
    LATIN CAPITAL LETTER NJ
<U006F> <o>; <BLANK>; <LOWERCASE>; <U006F> %
    LATIN SMALL LETTER O
<U004F> <o>; <BLANK>; <UPPERCASE>; <U004F> %
    LATIN CAPITAL LETTER O
<U00F3> <o>; <ACUTE>; <LOWERCASE>; <U00F3> %
    LATIN SMALL LETTER O WITH ACUTE
<U00D3> <o>; <ACUTE>; <UPPERCASE>; <U00D3> %
    LATIN CAPITAL LETTER O WITH ACUTE
<U00F2> <o>; <GRAVE>; <LOWERCASE>; <U00F2> %
    LATIN SMALL LETTER O WITH GRAVE
<U00D2> <o>; <GRAVE>; <UPPERCASE>; <U00D2> %
    LATIN CAPITAL LETTER O WITH GRAVE
<U014F> <o>; <BREVE>; <LOWERCASE>; <U014F> %
    LATIN SMALL LETTER O WITH BREVE
<U014E> <o>; <BREVE>; <UPPERCASE>; <U014E> %
    LATIN CAPITAL LETTER O WITH BREVE
<U00F4> <o>; <CIRCUMFLEX>; <LOWERCASE>;
    <U00F4> % LATIN SMALL LETTER O WITH
    CIRCUMFLEX
<U00D4> <o>; <CIRCUMFLEX>; <UPPERCASE>;
    <U00D4> % LATIN CAPITAL LETTER O WITH
    CIRCUMFLEX
<U1ED1> <o>; <CIRCUMFLEX><ACUTE>;
    <LOWERCASE>; <U1ED1> % LATIN SMALL LETTER
    O WITH CIRCUMFLEX AND ACUTE
<U1ED0> <o>; <CIRCUMFLEX><ACUTE>;
    <UPPERCASE>; <U1ED0> % LATIN CAPITAL
    LETTER O WITH CIRCUMFLEX AND ACUTE
<U1ED3> <o>; <CIRCUMFLEX><GRAVE>;
    <LOWERCASE>; <U1ED3> % LATIN SMALL LETTER
    O WITH CIRCUMFLEX AND GRAVE
<U1ED2> <o>; <CIRCUMFLEX><GRAVE>;
    <UPPERCASE>; <U1ED2> % LATIN CAPITAL
    LETTER O WITH CIRCUMFLEX AND GRAVE
<U1ED5> <o>; <CIRCUMFLEX><HOOK_ABOVE>;
    <LOWERCASE>; <U1ED5> % LATIN SMALL LETTER
    O WITH CIRCUMFLEX AND HOOK ABOVE
<U1ED4> <o>; <CIRCUMFLEX><HOOK_ABOVE>;
    <UPPERCASE>; <U1ED4> % LATIN CAPITAL
    LETTER O WITH CIRCUMFLEX AND HOOK ABOVE
<U1ED7> <o>; <CIRCUMFLEX><TILDE>;
    <LOWERCASE>; <U1ED7> % LATIN SMALL LETTER
    O WITH CIRCUMFLEX AND TILDE
<U1ED6> <o>; <CIRCUMFLEX><TILDE>;
    <UPPERCASE>; <U1ED6> % LATIN CAPITAL
    LETTER O WITH CIRCUMFLEX AND TILDE
<U1ED9> <o>; <CIRCUMFLEX><DOT_BELOW>;
    <LOWERCASE>; <U1ED9> % LATIN SMALL LETTER
    O WITH CIRCUMFLEX AND DOT BELOW
<U1ED8> <o>; <CIRCUMFLEX><DOT_BELOW>;
    <UPPERCASE>; <U1ED8> % LATIN CAPITAL
    LETTER O WITH CIRCUMFLEX AND DOT BELOW
<U01D2> <o>; <CARON>; <LOWERCASE>; <U01D2> %
    LATIN SMALL LETTER O WITH CARON
<U01D1> <o>; <CARON>; <UPPERCASE>; <U01D1> %
    LATIN CAPITAL LETTER O WITH CARON
<U00F6> <o>; <DIAERESIS>; <LOWERCASE>;
    <U00F6> % LATIN SMALL LETTER O WITH
    DIAERESIS
<U00D6> <o>; <DIAERESIS>; <UPPERCASE>;
    <U00D6> % LATIN CAPITAL LETTER O WITH
    DIAERESIS
<U0151> <o>; <DOUBLE_ACUTE>; <LOWERCASE>;
    <U0151> % LATIN SMALL LETTER O WITH
    DOUBLE ACUTE
<U0150> <o>; <DOUBLE_ACUTE>; <UPPERCASE>;
    <U0150> % LATIN CAPITAL LETTER O WITH
    DOUBLE ACUTE
<U1ECF> <o>; <HOOK_ABOVE>; <LOWERCASE>;
    <U1ECF> % LATIN SMALL LETTER O WITH HOOK
    ABOVE
<U1ECE> <o>; <HOOK_ABOVE>; <UPPERCASE>;
    <U1ECE> % LATIN CAPITAL LETTER O WITH
    HOOK ABOVE
<U00F5> <o>; <TILDE>; <LOWERCASE>; <U00F5> %
    LATIN SMALL LETTER O WITH TILDE

<U00D5> <o>; <TILDE>; <UPPERCASE>; <U00D5> %
    LATIN CAPITAL LETTER O WITH TILDE
<U1ECD> <o>; <DOT_BELOW>; <LOWERCASE>;
    <U1ECD> % LATIN SMALL LETTER O WITH DOT
    BELOW
<U1ECC> <o>; <DOT_BELOW>; <UPPERCASE>;
    <U1ECC> % LATIN CAPITAL LETTER O WITH DOT
    BELOW
<U01EB> <o>; <OGONEK>; <LOWERCASE>; <U01EB> %
    LATIN SMALL LETTER O WITH OGONEK
<U01EA> <o>; <OGONEK>; <UPPERCASE>; <U01EA> %
    LATIN CAPITAL LETTER O WITH OGONEK
<U014D> <o>; <MACRON>; <LOWERCASE>; <U014D> %
    LATIN SMALL LETTER O WITH MACRON
<U014C> <o>; <MACRON>; <UPPERCASE>; <U014C> %
    LATIN CAPITAL LETTER O WITH MACRON
<U01A1> <o>; <HORN>; <LOWERCASE>; <U01A1> %
    LATIN SMALL LETTER O WITH HORN
<U01A0> <o>; <HORN>; <UPPERCASE>; <U01A0> %
    LATIN CAPITAL LETTER O WITH HORN
<U1EDB> <o>; <HORN><ACUTE>; <LOWERCASE>;
    <U1EDB> % LATIN SMALL LETTER O WITH HORN
    AND ACUTE
<U1EDA> <o>; <HORN><ACUTE>; <UPPERCASE>;
    <U1EDA> % LATIN CAPITAL LETTER O WITH
    HORN AND ACUTE
<U1EDD> <o>; <HORN><GRAVE>; <LOWERCASE>;
    <U1EDD> % LATIN SMALL LETTER O WITH HORN
    AND GRAVE
<U1EDC> <o>; <HORN><GRAVE>; <UPPERCASE>;
    <U1EDC> % LATIN CAPITAL LETTER O WITH
    HORN AND GRAVE
<U1EDF> <o>; <HORN><HOOK_ABOVE>; <LOWERCASE>;
    <U1EDF> % LATIN SMALL LETTER O WITH HORN
    AND HOOK ABOVE
<U1EDE> <o>; <HORN><HOOK_ABOVE>; <UPPERCASE>;
    <U1EDE> % LATIN CAPITAL LETTER O WITH
    HORN AND HOOK ABOVE
<U1EE1> <o>; <HORN><TILDE>; <LOWERCASE>;
    <U1EE1> % LATIN SMALL LETTER O WITH HORN
    AND TILDE
<U1EE0> <o>; <HORN><TILDE>; <UPPERCASE>;
    <U1EE0> % LATIN CAPITAL LETTER O WITH
    HORN AND TILDE
<U1EE3> <o>; <HORN><DOT_BELOW>; <LOWERCASE>;
    <U1EE3> % LATIN SMALL LETTER O WITH HORN
    AND DOT BELOW
<U1EE2> <o>; <HORN><DOT_BELOW>; <UPPERCASE>;
    <U1EE2> % LATIN CAPITAL LETTER O WITH
    HORN AND DOT BELOW
<U00F8> <o>; <MODIFIED1>; <LOWERCASE>;
    <U00F8> % LATIN SMALL LETTER O WITH
    STROKE
<U00D8> <o>; <MODIFIED1>; <UPPERCASE>;
    <U00D8> % LATIN CAPITAL LETTER O WITH
    STROKE
<U01FF> <o>; <MODIFIED1><ACUTE>; <LOWERCASE>;
    <U01FF> % LATIN SMALL LETTER O WITH
    STROKE AND ACUTE
<U01FE> <o>; <MODIFIED1><ACUTE>; <UPPERCASE>;
    <U01FE> % LATIN CAPITAL LETTER O WITH
    STROKE AND ACUTE
<U0153> "<o><e>"; <MODIFIED2><MODIFIED2>;
    <LOWERCASE><LOWERCASE>; <U006F><U0065> %
    LATIN SMALL LIGATURE OE
<U0153> "<o><e>"; <MODIFIED2><MODIFIED2>;
    <UPPERCASE><UPPERCASE>; <U004F><U0045> %
    LATIN CAPITAL LIGATURE OE
<U0070> <p>; <BLANK>; <LOWERCASE>; <U0070> %
    LATIN SMALL LETTER P
<U0050> <p>; <BLANK>; <UPPERCASE>; <U0050> %
    LATIN CAPITAL LETTER P
<U1E55> <p>; <ACUTE>; <LOWERCASE>; <U1E55> %
    LATIN SMALL LETTER P WITH ACUTE
<U1E54> <p>; <ACUTE>; <UPPERCASE>; <U1E54> %
    LATIN CAPITAL LETTER P WITH ACUTE
<U1E57> <p>; <DOT_ABOVE>; <LOWERCASE>;
    <U1E57> % LATIN SMALL LETTER P WITH DOT
    ABOVE
<U1E56> <p>; <DOT_ABOVE>; <UPPERCASE>;
    <U1E56> % LATIN CAPITAL LETTER P WITH DOT
    ABOVE
<U0071> <q>; <BLANK>; <LOWERCASE>; <U0071> %
    LATIN SMALL LETTER Q
<U0051> <q>; <BLANK>; <UPPERCASE>; <U0051> %
    LATIN CAPITAL LETTER Q
```

<U0072> <r>; <BLANK>; <LOWERCASE>; <U0072> %
LATIN SMALL LETTER R
<U0052> <r>; <BLANK>; <UPPERCASE>; <U0052> %
LATIN CAPITAL LETTER R
<U0155> <r>; <ACUTE>; <LOWERCASE>; <U0155> %
LATIN SMALL LETTER R WITH ACUTE
<U0154> <r>; <ACUTE>; <UPPERCASE>; <U0154> %
LATIN CAPITAL LETTER R WITH ACUTE
<U0159> <r>; <CARON>; <LOWERCASE>; <U0159> %
LATIN SMALL LETTER R WITH CARON
<U0158> <r>; <CARON>; <UPPERCASE>; <U0158> %
LATIN CAPITAL LETTER R WITH CARON
<U1E59> <r>; <DOT_ABOVE>; <LOWERCASE>;
    <U1E59> % LATIN SMALL LETTER R WITH DOT
    ABOVE
<U1E58> <r>; <DOT_ABOVE>; <UPPERCASE>;
    <U1E58> % LATIN CAPITAL LETTER R WITH DOT
    ABOVE
<U1E5B> <r>; <DOT_BELOW>; <LOWERCASE>;
    <U1E5B> % LATIN SMALL LETTER R WITH DOT
    BELOW
<U1E5A> <r>; <DOT_BELOW>; <UPPERCASE>;
    <U1E5A> % LATIN CAPITAL LETTER R WITH DOT
    BELOW
<U0157> <r>; <CEDILLA>; <LOWERCASE>; <U0157>
    % LATIN SMALL LETTER R WITH CEDILLA
<U0156> <r>; <CEDILLA>; <UPPERCASE>; <U0156>
    % LATIN CAPITAL LETTER R WITH CEDILLA
<U0073> <s>; <BLANK>; <LOWERCASE>; <U0073> %
LATIN SMALL LETTER S
<U0053> <s>; <BLANK>; <UPPERCASE>; <U0053> %
LATIN CAPITAL LETTER S
<U015B> <s>; <ACUTE>; <LOWERCASE>; <U015B> %
LATIN SMALL LETTER S WITH ACUTE
<U015A> <s>; <ACUTE>; <UPPERCASE>; <U015A> %
LATIN CAPITAL LETTER S WITH ACUTE
<U015D> <s>; <CIRCUMFLEX>; <LOWERCASE>;
    <U015D> % LATIN SMALL LETTER S WITH
    CIRCUMFLEX
<U015C> <s>; <CIRCUMFLEX>; <UPPERCASE>;
    <U015C> % LATIN CAPITAL LETTER S WITH
    CIRCUMFLEX
<U0161> <s>; <CARON>; <LOWERCASE>; <U0161> %
LATIN SMALL LETTER S WITH CARON
<U0160> <s>; <CARON>; <UPPERCASE>; <U0160> %
LATIN CAPITAL LETTER S WITH CARON
<U1E61> <s>; <DOT_ABOVE>; <LOWERCASE>;
    <U1E61> % LATIN SMALL LETTER S WITH DOT
    ABOVE
<U1E60> <s>; <DOT_ABOVE>; <UPPERCASE>;
    <U1E60> % LATIN CAPITAL LETTER S WITH DOT
    ABOVE
<U1E63> <s>; <DOT_BELOW>; <LOWERCASE>;
    <U1E63> % LATIN SMALL LETTER S WITH DOT
    BELOW
<U1E62> <s>; <DOT_BELOW>; <UPPERCASE>;
    <U1E62> % LATIN CAPITAL LETTER S WITH DOT
    BELOW
<U015F> <s>; <CEDILLA>; <LOWERCASE>; <U015F>
    % LATIN SMALL LETTER S WITH CEDILLA
<U015E> <s>; <CEDILLA>; <UPPERCASE>; <U015E>
    % LATIN CAPITAL LETTER S WITH CEDILLA
<U00DF> "<s><s>"; <MODIFIED1><BLANK>;
    <LOWERCASE><LOWERCASE>; <U017F><U0073>
    LATIN SMALL LETTER SHARP S
<UFB06> "<s><t>"; <BLANK><BLANK>;
    <LOWERCASE><LOWERCASE>; <U0073><U0074> %
    LATIN SMALL LIGATURE ST
<U0074> <t>; <BLANK>; <LOWERCASE>; <U0074> %
LATIN SMALL LETTER T
<U0054> <t>; <BLANK>; <UPPERCASE>; <U0054> %
LATIN CAPITAL LETTER T
<U1E71> <t>; <CIRCUMFLEX_BELOW>; <LOWERCASE>;
    <U1E71> % LATIN SMALL LETTER T WITH
    CIRCUMFLEX BELOW
<U1E70> <t>; <CIRCUMFLEX_BELOW>; <UPPERCASE>;
    <U1E70> % LATIN CAPITAL LETTER T WITH
    CIRCUMFLEX BELOW
<U0165> <t>; <CARON>; <LOWERCASE>; <U0165> %
LATIN SMALL LETTER T WITH CARON
<U0164> <t>; <CARON>; <UPPERCASE>; <U0164> %
LATIN CAPITAL LETTER T WITH CARON
<U1E97> <t>; <DIAERESIS>; <LOWERCASE>;
    <U1E97> % LATIN SMALL LETTER T WITH
    DIAERESIS

<U1E6B> <t>; <DOT_ABOVE>; <LOWERCASE>;
    <U1E6B> % LATIN SMALL LETTER T WITH DOT
    ABOVE
<U1E6A> <t>; <DOT_ABOVE>; <UPPERCASE>;
    <U1E6A> % LATIN CAPITAL LETTER T WITH DOT
    ABOVE
<U1E6D> <t>; <DOT_BELOW>; <LOWERCASE>;
    <U1E6D> % LATIN SMALL LETTER T WITH DOT
    BELOW
<U1E6C> <t>; <DOT_BELOW>; <UPPERCASE>;
    <U1E6C> % LATIN CAPITAL LETTER T WITH DOT
    BELOW
<U0163> <t>; <CEDILLA>; <LOWERCASE>; <U0163>
    % LATIN SMALL LETTER T WITH CEDILLA
<U0162> <t>; <CEDILLA>; <UPPERCASE>; <U0162>
    % LATIN CAPITAL LETTER T WITH CEDILLA
<U1E6F> <t>; <MACRON_BELOW>; <LOWERCASE>;
    <U1E6F> % LATIN SMALL LETTER T WITH LINE
    BELOW
<U1E6E> <t>; <MACRON_BELOW>; <UPPERCASE>;
    <U1E6E> % LATIN CAPITAL LETTER T WITH
    LINE BELOW
<U0167> <t>; <MODIFIED1>; <LOWERCASE>;
    <U0167> % LATIN SMALL LETTER T WITH
    STROKE
<U0166> <t>; <MODIFIED1>; <UPPERCASE>;
    <U0166> % LATIN CAPITAL LETTER T WITH
    STROKE
<U02A6> "<t><s>"; <BLANK><BLANK>;
    <LOWERCASE><LOWERCASE>; <U0074><U0073> %
    LATIN SMALL LETTER TS DIGRAPH
<U0075> <u>; <BLANK>; <LOWERCASE>; <U0075> %
LATIN SMALL LETTER U
<U0055> <u>; <BLANK>; <UPPERCASE>; <U0055> %
LATIN CAPITAL LETTER U
<U00FA> <u>; <ACUTE>; <LOWERCASE>; <U00FA> %
LATIN SMALL LETTER U WITH ACUTE
<U00DA> <u>; <ACUTE>; <UPPERCASE>; <U00DA> %
LATIN CAPITAL LETTER U WITH ACUTE
<U00F9> <u>; <GRAVE>; <LOWERCASE>; <U00F9> %
LATIN SMALL LETTER U WITH GRAVE
<U00D9> <u>; <GRAVE>; <UPPERCASE>; <U00D9> %
LATIN CAPITAL LETTER U WITH GRAVE
<U016D> <u>; <BREVE>; <LOWERCASE>; <U016D> %
LATIN SMALL LETTER U WITH BREVE
<U016C> <u>; <BREVE>; <UPPERCASE>; <U016C> %
LATIN CAPITAL LETTER U WITH BREVE
<U00FB> <u>; <CIRCUMFLEX>; <LOWERCASE>;
    <U00FB> % LATIN SMALL LETTER U WITH
    CIRCUMFLEX
<U00DB> <u>; <CIRCUMFLEX>; <UPPERCASE>;
    <U00DB> % LATIN CAPITAL LETTER U WITH
    CIRCUMFLEX
<U01D4> <u>; <CARON>; <LOWERCASE>; <U01D4> %
LATIN SMALL LETTER U WITH CARON
<U01D3> <u>; <CARON>; <UPPERCASE>; <U01D3> %
LATIN CAPITAL LETTER U WITH CARON
<U016F> <u>; <RING_ABOVE>; <LOWERCASE>;
    <U016F> % LATIN SMALL LETTER U WITH RING
    ABOVE
<U016E> <u>; <RING_ABOVE>; <UPPERCASE>;
    <U016E> % LATIN CAPITAL LETTER U WITH
    RING ABOVE
<U00FC> <u>; <DIAERESIS>; <LOWERCASE>;
    <U00FC> % LATIN SMALL LETTER U WITH
    DIAERESIS
<U00DC> <u>; <DIAERESIS>; <UPPERCASE>;
    <U00DC> % LATIN CAPITAL LETTER U WITH
    DIAERESIS
<U0171> <u>; <DOUBLE_ACUTE>; <LOWERCASE>;
    <U0171> % LATIN SMALL LETTER U WITH
    DOUBLE ACUTE
<U0170> <u>; <DOUBLE_ACUTE>; <UPPERCASE>;
    <U0170> % LATIN CAPITAL LETTER U WITH
    DOUBLE ACUTE
<U1EE7> <u>; <HOOK_ABOVE>; <LOWERCASE>;
    <U1EE7> % LATIN SMALL LETTER U WITH HOOK
    ABOVE
<U1EE6> <u>; <HOOK_ABOVE>; <UPPERCASE>;
    <U1EE6> % LATIN CAPITAL LETTER U WITH
    HOOK ABOVE
<U0169> <u>; <TILDE>; <LOWERCASE>; <U0169> %
LATIN SMALL LETTER U WITH TILDE
<U0168> <u>; <TILDE>; <UPPERCASE>; <U0168> %
LATIN CAPITAL LETTER U WITH TILDE

<U1EE5> <u>; <DOT_BELOW>; <LOWERCASE>;
    <U1EE5> % LATIN SMALL LETTER U WITH DOT
    BELOW
<U1EE4> <u>; <DOT_BELOW>; <UPPERCASE>;
    <U1EE4> % LATIN CAPITAL LETTER U WITH DOT
    BELOW
<U0173> <u>; <OGONEK>; <LOWERCASE>; <U0173> %
LATIN SMALL LETTER U WITH OGONEK
<U0172> <u>; <OGONEK>; <UPPERCASE>; <U0172> %
LATIN CAPITAL LETTER U WITH OGONEK
<U016B> <u>; <MACRON>; <LOWERCASE>; <U016B> %
LATIN SMALL LETTER U WITH MACRON
<U016A> <u>; <MACRON>; <UPPERCASE>; <U016A> %
LATIN CAPITAL LETTER U WITH MACRON
<U01B0> <u>; <HORN>; <LOWERCASE>; <U01B0> %
LATIN SMALL LETTER U WITH HORN
<U01AF> <u>; <HORN>; <UPPERCASE>; <U01AF> %
LATIN CAPITAL LETTER U WITH HORN
<U1EE9> <u>; <HORN><ACUTE>; <LOWERCASE>;
    <U1EE9> % LATIN SMALL LETTER U WITH HORN
    AND ACUTE
<U1EE8> <u>; <HORN><ACUTE>; <UPPERCASE>;
    <U1EE8> % LATIN CAPITAL LETTER U WITH
    HORN AND ACUTE
<U1EEB> <u>; <HORN><GRAVE>; <LOWERCASE>;
    <U1EEB> % LATIN SMALL LETTER U WITH HORN
    AND GRAVE
<U1EEA> <u>; <HORN><GRAVE>; <UPPERCASE>;
    <U1EEA> % LATIN CAPITAL LETTER U WITH
    HORN AND GRAVE
<U1EED> <u>; <HORN><HOOK_ABOVE>; <LOWERCASE>;
    <U1EED> % LATIN SMALL LETTER U WITH HORN
    AND HOOK ABOVE
<U1EEC> <u>; <HORN><HOOK_ABOVE>; <UPPERCASE>;
    <U1EEC> % LATIN CAPITAL LETTER U WITH
    HORN AND HOOK ABOVE
<U1EEF> <u>; <HORN><TILDE>; <LOWERCASE>;
    <U1EEF> % LATIN SMALL LETTER U WITH HORN
    AND TILDE
<U1EEE> <u>; <HORN><TILDE>; <UPPERCASE>;
    <U1EEE> % LATIN CAPITAL LETTER U WITH
    HORN AND TILDE
<U1EF1> <u>; <HORN><DOT_BELOW>; <LOWERCASE>;
    <U1EF1> % LATIN SMALL LETTER U WITH HORN
    AND DOT BELOW
<U1EF0> <u>; <HORN><DOT_BELOW>; <UPPERCASE>;
    <U1EF0> % LATIN CAPITAL LETTER U WITH
    HORN AND DOT BELOW
<U0076> <v>; <BLANK>; <LOWERCASE>; <U0076> %
LATIN SMALL LETTER V
<U0056> <v>; <BLANK>; <UPPERCASE>; <U0056> %
LATIN CAPITAL LETTER V
<U1E7D> <v>; <TILDE>; <LOWERCASE>; <U1E7D> %
LATIN SMALL LETTER V WITH TILDE
<U1E7C> <v>; <TILDE>; <UPPERCASE>; <U1E7C> %
LATIN CAPITAL LETTER V WITH TILDE
<U1E7F> <v>; <DOT_BELOW>; <LOWERCASE>;
    <U1E7F> % LATIN SMALL LETTER V WITH DOT
    BELOW
<U1E7E> <v>; <DOT_BELOW>; <UPPERCASE>;
    <U1E7E> % LATIN CAPITAL LETTER V WITH DOT
    BELOW
<U0077> <w>; <BLANK>; <LOWERCASE>; <U0077> %
LATIN SMALL LETTER W
<U0057> <w>; <BLANK>; <UPPERCASE>; <U0057> %
LATIN CAPITAL LETTER W
<U1E83> <w>; <ACUTE>; <LOWERCASE>; <U1E83> %
LATIN SMALL LETTER W WITH ACUTE
<U1E82> <w>; <ACUTE>; <UPPERCASE>; <U1E82> %
LATIN CAPITAL LETTER W WITH ACUTE
<U1E81> <w>; <GRAVE>; <LOWERCASE>; <U1E81> %
LATIN SMALL LETTER W WITH GRAVE
<U1E80> <w>; <GRAVE>; <UPPERCASE>; <U1E80> %
LATIN CAPITAL LETTER W WITH GRAVE
<U0175> <w>; <CIRCUMFLEX>; <LOWERCASE>;
    <U0175> % LATIN SMALL LETTER W WITH
    CIRCUMFLEX
<U0174> <w>; <CIRCUMFLEX>; <UPPERCASE>;
    <U0174> % LATIN CAPITAL LETTER W WITH
    CIRCUMFLEX
<U1E85> <w>; <DIAERESIS>; <LOWERCASE>;
    <U1E85> % LATIN SMALL LETTER W WITH
    DIAERESIS
<U1E84> <w>; <DIAERESIS>; <UPPERCASE>;
    <U1E84> % LATIN CAPITAL LETTER W WITH
    DIAERESIS

<U1E87> <w>; <DOT_ABOVE>; <LOWERCASE>;
    <U1E87> % LATIN SMALL LETTER W WITH DOT
    ABOVE
<U1E86> <w>; <DOT_ABOVE>; <UPPERCASE>;
    <U1E86> % LATIN CAPITAL LETTER W WITH DOT
    ABOVE
<U1E89> <w>; <DOT_BELOW>; <LOWERCASE>;
    <U1E89> % LATIN SMALL LETTER W WITH DOT
    BELOW
<U1E88> <w>; <DOT_BELOW>; <UPPERCASE>;
    <U1E88> % LATIN CAPITAL LETTER W WITH DOT
    BELOW
<U0078> <x>; <BLANK>; <LOWERCASE>; <U0078> %
    LATIN SMALL LETTER X
<U0058> <x>; <BLANK>; <UPPERCASE>; <U0058> %
    LATIN CAPITAL LETTER X
<U1E8D> <x>; <DIAERESIS>; <LOWERCASE>;
    <U1E8D> % LATIN SMALL LETTER X WITH
    DIAERESIS
<U1E8C> <x>; <DIAERESIS>; <UPPERCASE>;
    <U1E8C> % LATIN CAPITAL LETTER X WITH
    DIAERESIS
<U1E8B> <x>; <DOT_ABOVE>; <LOWERCASE>;
    <U1E8B> % LATIN SMALL LETTER X WITH DOT
    ABOVE
<U1E8A> <x>; <DOT_ABOVE>; <UPPERCASE>;
    <U1E8A> % LATIN CAPITAL LETTER X WITH DOT
    ABOVE
<U0079> <y>; <BLANK>; <LOWERCASE>; <U0079> %
    LATIN SMALL LETTER Y
<U0059> <y>; <BLANK>; <UPPERCASE>; <U0059> %
    LATIN CAPITAL LETTER Y
<U00FD> <y>; <ACUTE>; <LOWERCASE>; <U00FD> %
    LATIN SMALL LETTER Y WITH ACUTE
<U00DD> <y>; <ACUTE>; <UPPERCASE>; <U00DD> %
    LATIN CAPITAL LETTER Y WITH ACUTE
<U1EF3> <y>; <GRAVE>; <LOWERCASE>; <U1EF3> %
    LATIN SMALL LETTER Y WITH GRAVE
<U1EF2> <y>; <GRAVE>; <UPPERCASE>; <U1EF2> %
    LATIN CAPITAL LETTER Y WITH GRAVE

<U0177> <y>; <CIRCUMFLEX>; <LOWERCASE>;
    <U0177> % LATIN SMALL LETTER Y WITH
    CIRCUMFLEX
<U0176> <y>; <CIRCUMFLEX>; <UPPERCASE>;
    <U0176> % LATIN CAPITAL LETTER Y WITH
    CIRCUMFLEX
<U00FF> <y>; <DIAERESIS>; <LOWERCASE>;
    <U00FF> % LATIN SMALL LETTER Y WITH
    DIAERESIS
<U0178> <y>; <DIAERESIS>; <UPPERCASE>;
    <U0178> % LATIN CAPITAL LETTER Y WITH
    DIAERESIS
<U1EF7> <y>; <HOOK_ABOVE>; <LOWERCASE>;
    <U1EF7> % LATIN SMALL LETTER Y WITH HOOK
    ABOVE
<U1EF6> <y>; <HOOK_ABOVE>; <UPPERCASE>;
    <U1EF6> % LATIN CAPITAL LETTER Y WITH
    HOOK ABOVE
<U1EF9> <y>; <TILDE>; <LOWERCASE>; <U1EF9> %
    LATIN SMALL LETTER Y WITH TILDE
<U1EF8> <y>; <TILDE>; <UPPERCASE>; <U1EF8> %
    LATIN CAPITAL LETTER Y WITH TILDE
<U1E8F> <y>; <DOT_ABOVE>; <LOWERCASE>;
    <U1E8F> % LATIN SMALL LETTER Y WITH DOT
    ABOVE
<U1E8E> <y>; <DOT_ABOVE>; <UPPERCASE>;
    <U1E8E> % LATIN CAPITAL LETTER Y WITH DOT
    ABOVE
<U1EF5> <y>; <DOT_BELOW>; <LOWERCASE>;
    <U1EF5> % LATIN SMALL LETTER Y WITH DOT
    BELOW
<U1EF4> <y>; <DOT_BELOW>; <UPPERCASE>;
    <U1EF4> % LATIN CAPITAL LETTER Y WITH DOT
    BELOW
<U007A> <z>; <BLANK>; <LOWERCASE>; <U007A> %
    LATIN SMALL LETTER Z
<U005A> <z>; <BLANK>; <UPPERCASE>; <U005A> %
    LATIN CAPITAL LETTER Z
<U017A> <z>; <ACUTE>; <LOWERCASE>; <U017A> %
    LATIN SMALL LETTER Z WITH ACUTE

<U0179> <z>; <ACUTE>; <UPPERCASE>; <U0179> %
    LATIN CAPITAL LETTER Z WITH ACUTE
<U1E91> <z>; <CIRCUMFLEX>; <LOWERCASE>;
    <U1E91> % LATIN SMALL LETTER Z WITH
    CIRCUMFLEX
<U1E90> <z>; <CIRCUMFLEX>; <UPPERCASE>;
    <U1E90> % LATIN CAPITAL LETTER Z WITH
    CIRCUMFLEX
<U017E> <z>; <CARON>; <LOWERCASE>; <U017E> %
    LATIN SMALL LETTER Z WITH CARON
<U017D> <z>; <CARON>; <UPPERCASE>; <U017D> %
    LATIN CAPITAL LETTER Z WITH CARON
<U017C> <z>; <DOT_ABOVE>; <LOWERCASE>;
    <U017C> % LATIN SMALL LETTER Z WITH DOT
    ABOVE
<U017B> <z>; <DOT_ABOVE>; <UPPERCASE>;
    <U017B> % LATIN CAPITAL LETTER Z WITH DOT
    ABOVE
<U1E93> <z>; <DOT_BELOW>; <LOWERCASE>;
    <U1E93> % LATIN SMALL LETTER Z WITH DOT
    BELOW
<U1E92> <z>; <DOT_BELOW>; <UPPERCASE>;
    <U1E92> % LATIN CAPITAL LETTER Z WITH DOT
    BELOW
<U00FE> <thorn>; <BLANK>; <LOWERCASE>;
    <U00FE> % LATIN SMALL LETTER THORN
<U00DE> <thorn>; <BLANK>; <UPPERCASE>;
    <U00DE> % LATIN CAPITAL LETTER THORN
%%
reorder section after
    ISO12199CollatingSymbols <U0000>
reorder section after ISO12199Digits <U0020>
reorder section after ISO12199LatinLetters
    <U2182>
%%
order start ISO12199Digits forward; forward;
    forward; forward
order start ISO12199LatinLetters forward;
    forward; forward; forward

Bibliography

[1] ISO 9:1995, *Information and documentation — Transliteration of Cyrillic characters into Latin characters — Slavic and non-Slavic languages.*

[2] ISO 233:1984, *Documentation — Transliteration of Arabic characters into Latin characters.*

[3] ISO 233-2:1993, *Information and documentation — Transliteration of Arabic characters into Latin characters — Part 2: Arabic language — Simplified transliteration.*

[4] ISO 233-3:1999, *Information and documentation — Transliteration of Arabic characters into Latin characters — Part 3: Persian language — Simplified transliteration.*

[5] ISO 259:1984, *Documentation — Transliteration of Hebrew characters into Latin characters.*

[6] ISO 259-2:1994, *Information and documentation — Transliteration of Hebrew characters into Latin characters — Part 2: Simplified transliteration.*

[7] ISO 259-3:—[3], *Information and documentation — Transliteration of Hebrew characters into Latin characters — Part 3: Phonemic conversion.*

[8] ISO 639:1988, *Code for the representation of names of languages.*

[9] ISO 639-2:1998, *Codes for the representation of names of languages — Part 2: Alpha-3 code.*

[10] ISO 843:1997, *Information and documentation — Conversion of Greek characters into Latin characters.*

[11] ISO 999:1996, *Information and documentation — Guidelines for the content, organization and presentation of indexes.*

[12] ISO/IEC 2022:1994, *Information technology — Character code structure and extension techniques.*

[13] ISO 5426:1983, *Extension of the Latin alphabet coded character set for bibliographic information interchange.*

[14] ISO 5426-2:1996, *Information and documentation — Extension of the Latin alphabet coded character set for bibliographic information interchange — Part 2: Latin characters used in minor European languages and obsolete typography.*

[15] ISO 5427:1984, *Extension of the Cyrillic alphabet coded character set for bibliographic information interchange.*

[16] ISO 5428:1984, *Greek alphabet coded character set for bibliographic information interchange.*

[17] ISO/IEC 6429:1992, *Information technology — Control functions for coded character sets.*

[18] ISO 6438:1983, *Documentation — African coded character set for bibliographic information interchange.*

[19] ISO/IEC 6937:1994, *Information technology — Coded graphic character sets for text communication — Latin alphabet.*

[20] ISO 7098:1991, *Information and documentation — Romanization of Chinese.*

[21] ISO/IEC 8859-1:1998, *Information technology — 8-bit single-byte coded graphic character sets — Part 1: Latin alphabet No. 1.*

---

3) To be published.

[22] ISO/IEC 8859-2:1999, *Information technology — 8-bit single-byte coded graphic character sets — Part 2: Latin alphabet No. 2.*

[23] ISO/IEC 8859-3:1999, *Information technology — 8-bit single-byte coded graphic character sets — Part 3: Latin alphabet No. 3.*

[24] ISO/IEC 8859-4:1998, *Information technology — 8-bit single-byte coded graphic character sets — Part 4: Latin alphabet No. 4.*

[25] ISO/IEC 8859-5:1999, *Information technology — 8-bit single-byte coded graphic character sets — Part 5: Latin/Cyrillic alphabet.*

[26] ISO/IEC 8859-6:1999, *Information technology — 8-bit single-byte coded graphic character sets — Part 6: Latin/Arabic alphabet.*

[27] ISO 8859-7:1987, *Information processing — 8-bit single-byte coded graphic character sets — Part 7: Latin/Greek alphabet.*

[28] ISO/IEC 8859-8:1999, *Information technology — 8-bit single-byte coded graphic character sets — Part 8: Latin/Hebrew alphabet.*

[29] ISO/IEC 8859-9:1999, *Information technology — 8-bit single-byte coded graphic character sets — Part 9: Latin alphabet No. 5.*

[30] ISO/IEC 8859-10:1998, *Information technology — 8-bit single-byte coded graphic character sets — Part 10: Latin alphabet No. 6.*

[31] ISO/IEC 8859-11:—[4], *Information technology — 8-bit single-byte coded graphic character sets — Part 11: Latin/Thai character set.*

[32] ISO/IEC 8859-13:1998, *Information technology — 8-bit single-byte coded graphic character sets — Part 13: Latin alphabet No. 7.*

[33] ISO/IEC 8859-14:1998, *Information technology — 8-bit single-byte coded graphic character sets — Part 14: Latin alphabet No. 8 (Celtic).*

[34] ISO/IEC 8859-15:1999, *Information technology — 8-bit single-byte coded graphic character sets — Part 15: Latin alphabet No. 9.*

[35] ISO 8879:1986, *Information processing — Text and office systems — Standard Generalized Markup Language (SGML).*

[36] ISO/IEC 10367:1991, *Information technology — Standardized coded graphic character sets for use in 8-bit codes.*

[37] BS 1749:1985, *Recommendations for alphabetical arrangement and the filing order of numbers and symbols.* [British Standard]

[38] DIN 5007:1991, *Ordnen von Schriftzeichenfolgen (ABC-Regeln) [Rules for alphabetical arrangement (ABC rules)].* [Deutsche Norm]

[39] DIN 5007-2:1996, *Ordnen von Schriftzeichenfolgen — Teil 2: Ansetzungsregeln für die alphabetische Ordnung von Namen (Rules for alphabetical ordering — Part 2: Presentation of names).* [Deutsche Norm]

[40] DS 377:1980, *Alfabetiseringsregler (Alphabetical arrangements).* [Dansk Standard]

[41] NS 4103:1972, *Alfabetiseringsregler (Alphabetical arrangement).* [Norsk Standard]

---

4) To be published.

[42] ON A 2725:1973, *Regeln für das Ordnen von Schriftzeichenfolgen (ABC-Regeln) [Rules for alphabetical arrangement (ABC rules)]*. [Österreichische Norm]

[43] SFS 4600:1986, *Aakkostaminen ja siihen liityvä ryhmittely (Order of characters and numerals)*. [Suomalainen standardi]

[44] TN 34, *Swedish alphanumeric sorting*. [Statskontoret, Teknisk Norm]

[45] GAVARE, ROLF, *Alphabetical Ordering in a Lexicological Perspective — A computational model for exacting alphabetization of dictionary and encyclopedic material*. Pages 63-102 in Studies in Computer-Aided Lexicology, Stockholm, Almqvist & Wiksell International, 1988.

ISO 12199:2000(E)

ICS 01.020

Price based on 40 pages